

EXACT AND NON-STIFF SAMPLING OF HIGHLY OSCILLATORY SYSTEMS: AN IMPLICIT MASS-MATRIX PENALIZATION APPROACH

PETR PLECHÁČ* AND MATHIAS ROUSSET†

Abstract. We propose and analyze an implicit mass-matrix penalization (IMMP) technique which enables efficient and exact sampling of the (Boltzmann/Gibbs) canonical distribution associated to Hamiltonian systems with fast degrees of freedom (fDOFs). The penalty parameters enable arbitrary tuning of the timescale for the selected fDOFs, and the method is interpreted as an interpolation between the exact Hamiltonian dynamics and the dynamics with infinitely slow fDOFs (equivalent to geometrically corrected rigid constraints). This property translates in the associated numerical methods into a tunable trade-off between stability and dynamical modification. The penalization is based on an extended Hamiltonian with artificial constraints associated with each fDOF. By construction, the resulting dynamics is statistically exact with respect to the canonical distribution in position variables.

The algorithms can be easily implemented with standard geometric integrators with algebraic constraints given by the expected fDOFs, and has no additional complexity in terms of enforcing the constraint and force evaluations. The method is demonstrated on a high dimensional system with non-convex interactions. Prescribing the macroscopic dynamical timescale, it is shown that the IMMP method increases the time-step stability region with a gain that grows linearly with the size of the system. The latter property, as well as consistency of the macroscopic dynamics of the IMMP method is proved rigorously for linear interactions. Finally, when a large stiffness parameter is introduced, the IMMP method can be tuned to be asymptotically stable, converging towards the heuristically expected Markovian effective dynamics on the slow manifold.

Key words. Hamiltonian systems, NVT ensemble, slow/fast systems, Langevin dynamics, constrained dynamics, hybrid Monte Carlo.

AMS subject classifications. 65C05, 65C20, 82B20, 82B80, 82-08

1. Introduction. This paper deals with stochastic numerical integration and sampling of Hamiltonian systems with multiple timescales. The main motivation is to develop numerical methods which sample *exactly* canonical distributions, while integrating the fastest degrees of freedom only statistically. Furthermore, one also seeks good approximation of dynamical behavior, at least at large temporal scales.

Hamiltonian systems with multiple timescales typically appear in molecular dynamics (MD) simulations, which have become, with the aid of increasing computational power, a standard tool in many fields of physics, chemistry and biology. However, extending the simulations to physically relevant time-scales remains a major challenge for various large molecular systems. Due to the complexity of implicit methods, the time scales reachable by standard numerical methods are usually limited by the rapid oscillations of some particular degrees of freedom. Since the sampling dynamics has to be integrated for long times, the time-step restriction associated with fast oscillations/short time scales in molecular systems contributes to the high computational cost of such methods. Yet the physical necessity of resolving the fast degrees of freedom in simulations is often ambiguous, and efficient treatment of the fast time scales has motivated new interest in developing numerical schemes for the integration of such stiff systems.

The problem of integrating stiff forces is relevant both for the direct numerical simulation of the Hamiltonian dynamics, and for the less restrictive problem of designing a process that samples the canonical ensemble. Sampling from the canonical distribution can be achieved by Markov chain Monte Carlo (MCMC) algorithms based on an a priori knowledge of possible moves, with a Metropolis-Hastings acceptance/rejection corrector (a historical reference is [30]). For complex molecular systems however, such global moves remain unknown in general, and sampling methods consists generically in using either an Hamiltonian dynamics with a thermostat (e.g., a Langevin process), or a drifted random walk (Brownian dynamics), corresponding to over-damped Hamiltonian dynamics (see [7] for a review and references on classical sampling methods). Brownian dynamics of systems with multiple time scales suffers from similar

*Department of Mathematics, University of Tennessee and Oak Ridge National Laboratory, Knoxville, TN 37996-1300, USA, (plechac@math.utk.edu).

†INRIA Lille Nord-Europe, Villeneuve d'Ascq, France, (mathias.rousset@inria.fr)

stability restrictions (see [20, 39] for some practical issues on Brownian dynamics simulation for Molecular Dynamics).

Broadly speaking, one may start by recognizing two approaches to the numerical treatment of stiff systems:

- (i) Semi-implicit, multi-step integrators and their variants (e.g., the textbooks [34, 14], or the review paper [10] and references therein, [20] for Brownian dynamics), which attempt to resolve microscopic highly oscillatory dynamical behavior.
- (ii) Methods with direct constraints, where the highly oscillatory degrees of freedom are constrained to their equilibrium value (e.g., [34, 28, 38, 19] and references therein).

In spite of their differences the common key feature of all these methods is to balance a trade-off between stability restrictions and implicit time-stepping form, or in other words, between the computational effort associated with small time steps, and the computational cost of solving implicit equations implied by the stiffness.

The method developed in this paper aims at giving an appropriate interpolation between exact dynamics and constrained dynamics considered in the second family of methods mentioned above. Although constrained dynamics remove, in principle, the stiffness of the associated numerical scheme, it introduces new difficulties and numerical problems. As an approximation to the original dynamics it modifies important features of the system; most importantly, the original statistical distribution. The principal goal of the proposed method is to replace direct constraints by implicit mass-matrix penalization (IMMP), detailed in Section 3, which integrates fDOFs, but with a tunable mass penalty. This approach guarantees that the canonical distribution is computed *exactly*, and that a freely tunable trade-off between dynamical modification and stability can be obtained.

The idea of adjusting mass tensors in order to slow down fast degrees of freedom goes back to [3]. In this paper, the author proposes to modify the mass tensor with respect to the Hessian of the potential energy function in order to confine the frequency spectrum to low frequencies only. Two natural drawbacks of this procedure arise from the costly computation of the second-order derivatives of the potential, and from the bias introduced when the adjusted mass-tensor is adapted during the dynamics. Such an approach seems inevitable when the fDOFs are unknown, but in many cases, the fast degrees of freedom are explicitly given by the structure of the system (e.g., co-valent and angle bonds in molecular chains). To our knowledge, mass tensor modification have been used in practical MD simulations by increasing the mass of some well-chosen (e.g., light) atoms [26, 29]. The aim of this paper is to propose a more systematic mass-tensor modification strategy.

The proposed method relies on the assumption that the system Hamiltonian is separable with quadratic kinetic energy

$$H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q), \quad (1.1)$$

and that the “fast” degrees of freedom (ξ_1, \dots, ξ_n) are explicitly defined smooth functions of the system position

$$q = (q_1, \dots, q_d) \mapsto (\xi_1(q), \dots, \xi_n(q)). \quad (1.2)$$

We emphasize that the knowledge of “fast forces” is not required, and the variables ξ can be chosen arbitrarily. If the fDOFs are not identified the method retains its approximation properties while not performing efficiently. The fDOFs are penalized with a mass-tensor modification given by

$$M_\nu(q) = M + \nu^2 \nabla_q \xi M_z \nabla_q^T \xi, \quad (1.3)$$

where ν denotes the penalty intensity, and M_z a “virtual” mass matrix associated with the fDOFs. The modification does not impact motions orthogonal to the fDOFs. The position dependence of the mass-

penalization introduces a geometric bias. This bias is corrected by introducing an effective potential

$$V_{\text{fix},\nu}(q) = \frac{1}{2\beta} \ln(\det(M_\nu(q))) , \quad (1.4)$$

which will turn out to be a $\frac{1}{\nu}$ -perturbation of the usual Fixman corrector (see [19]) associated with the sub-manifolds defined by constraining the fDOFs ξ . The key point is then to use an implicit representation of the mass penalty with the aid of the extended Hamiltonian

$$\begin{cases} H_{\text{IMMP}}(p, p_z, q, z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + V(q) + V_{\text{fix},\nu}(q) , \\ \xi(q) = \frac{z}{\nu} . \end{cases} \quad (C_\nu) \quad (1.5)$$

The auxiliary degrees of freedom z are endowed with the “virtual” mass-matrix M_z . The constraints (C_ν) are applied in order to identify the auxiliary variables and the fDOFs ξ with a coupling intensity tuned by ν . The typical time scale of the fDOFs is thus enforced by the penalty ν . The system is coupled to a thermostat through a Langevin equation (2.3), which yields a stochastically perturbed dynamics that samples the equilibrium canonical distribution. We then obtain the following desirable properties:

1. The associated canonical equilibrium distribution in position is independent of the penalty ν .
2. The limit of vanishing penalization ($\nu = 0$) is the original full dynamics, enabling the construction of dynamically consistent numerical schemes.
3. The limit of infinite penalization is a standard effective constrained dynamics on the “slow” manifold associated with stiff constraints on ξ .
4. Numerical integrators can be obtained through a simple modification of standard integrators for effective dynamics with constraints, with equivalent computational complexity.

The dynamics associated with the IMMP Hamiltonian (1.5) is detailed in (3.6), see Section 3. Its numerical discretization by a leapfrog/Verlet scheme with constraints (usually called “RATTLE”) is given by (4.2). The flow can then be corrected by a Metropolis step to obtain exact canonical sampling, which is reminiscent of so-called Hybrid Monte-Carlo methods (see references in Section 4). When this correction is introduced, the gradient of the Fixman potential (1.4) need not be computed. The numerical aspects of the method are presented in Section 4.

By including a penalty, the proposed method modifies the original Hamiltonian. However, the mass penalty can be also thought of as depending on the time step $\nu := \nu(\delta t)$ leading to consistent schemes. In Section 6 and in Section 7, the penalty (and the fastest timescales) will even grow to infinity $\nu \rightarrow +\infty$ as the dimension of the system $N \rightarrow \infty$ or a stiffness parameter $\epsilon \rightarrow 0$. In both cases, we encounter generic situations where the IMMP method leads to dynamically consistent limits, a partial differential equation, and an effective dynamics on a “slow” manifold, respectively.

High-dimensional systems usually contain a large variety of timescales, and are therefore challenging test cases. A linear chain of particles with repulsive *non-convex* interactions is numerically studied in Section 5. It is shown that a large mass-penalty of order $\nu := \bar{\nu}N$, where N is the system size, induces a gain in time stepping of order N , while numerical evidence are given that the macroscopic dynamics, given by a formal stochastic partial differential equation, remains of order 1 (in particular, the convergence to equilibrium). Rigorous proofs with explicit scalings of these behaviors is provided for the linear case (i.e. harmonic interactions) in Section 6, and consistence of the IMMP macroscopic dynamics when the scaled penalty vanishes $\bar{\nu} \rightarrow 0$ is even demonstrated.

In Section 7, we introduce the stiffness parameter ϵ and show that the penalty intensity can be scaled as $\nu := \bar{\nu}/\epsilon$ in order to obtain asymptotically stable dynamics in the limit $\epsilon \rightarrow 0$. We prove that the dynamics converge towards the expected Markovian effective dynamics on the slow manifold. We also present analysis of stability properties of the proposed scheme.

Acknowledgments: The research of M.R. was partially supported by the EPSRC grant GR/S70883/01 while he was visiting Mathematics Institute, University of Warwick. The research of P.P. was partially

supported by the Office of Advanced Scientific Computing Research, U.S. Department of Energy; the work was partly done at the ORNL, which is managed by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725.

2. Formulation. We consider a Hamiltonian system in the phase-space $\mathbb{R}^d \times \mathbb{R}^d$ with the Hamiltonian H in the form

$$H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q), \quad (2.1)$$

We use generic matrix notation, for instance the Euclidean scalar product of two vectors $p_1, p_2 \in \mathbb{R}^N$ denoted by $p_1^T p_2$, and the gradients of mappings from \mathbb{R}^d to \mathbb{R}^n with respect to standard bases are represented by matrices

$$(\nabla_q^T \xi)_{ij} = (\nabla_q \xi)_{ji} = \frac{\partial \xi_i}{\partial q_j}, \quad i = 1, \dots, n, \quad j = 1, \dots, d.$$

When the system is thermostatted, i.e., kept at the constant temperature, the long time distribution of the system in the phase-space is given by the canonical equilibrium measure at the inverse temperature β (also called the NVT distribution) given by

$$\mu(dp dq) = \frac{1}{Z} e^{-\beta H(p, q)} dp dq, \quad Z = \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-\beta H(p, q)} dp dq, \quad (2.2)$$

with the normalization constant $Z < \infty$. The standard dynamics used to model thermostatted systems are given by Langevin processes.

DEFINITION 2.1 (Langevin process). *A Langevin process at the inverse temperature β with the Hamiltonian $H(q, p)$, $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$, the $d \times d$ dissipation matrix γ , and the fluctuation matrix σ is given by the stochastic differential equations*

$$\begin{cases} \dot{q} = \nabla_p H \\ \dot{p} = -\nabla_q H - \gamma \dot{q} + \sigma \dot{W} \end{cases}, \quad (2.3)$$

where \dot{W}^1 is a standard white noise (Wiener process), and $\sigma \in \mathbb{R}^d \times \mathbb{R}^d$ satisfies the fluctuation-dissipation identity

$$\sigma \sigma^T = \frac{2}{\beta} \gamma.$$

For any γ , the process is reversible with respect to the stationary canonical distribution (2.2). Furthermore, if γ is strictly positive definite, the process is ergodic. Throughout this paper, the usual global Lipschitz conditions (see [31]) on H and ξ are assumed, ensuring well-posedness of the considered stochastic differential equations. The analysis presented in the paper can be generalized to a position dependent dissipation matrix $\gamma = \gamma(q)$.

The mapping $\xi : \mathbb{R}^d \rightarrow \mathbb{R}^n$, defines $n \leq d$ degrees of freedom, given by smooth functions taking values in a neighborhood of 0. We assume that the mapping ξ is regular (i.e., with a non-degenerate Jacobian) in an open δ -neighborhood $\mathcal{O}_\delta = \{q \mid \|\xi(q)\| < \delta\}$ of $\xi^{-1}(0)$, hence defining a smooth sub-manifolds of \mathbb{R}^d denoted $\mathcal{M}_z = \xi^{-1}(z)$ for z in a neighborhood of 0. The dependence of the potential with respect to the degrees of freedom ξ is expected to be “stiff” in the second variable. In Section 7 we will introduce the

¹Throughout the paper, stochastic integrands have finite variation, so that stochastic integration (e.g. Ito or Stratonovitch) need not be specified.

stiffness parameter ϵ . We will assume in this section that such parameter dependence can be explicitly identified, and that the potential energy V can be written in the form

$$V(q) = U(q, \frac{\xi(q)}{\epsilon}), \quad (2.4)$$

where the function $U : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the coercivity condition $\lim_{z \rightarrow +\infty} U(q, z) = +\infty$. The fast degrees of freedom ξ of states at a given energy then remain in a closed neighborhood of the origin as the stiffness parameter $\epsilon \rightarrow 0$. In this limit the system is confined to the sub-manifold \mathcal{M}_0 , which is usually called the “slow manifold”.

3. The implicit mass-matrix penalization method. In this section we focus on properties of the IMMP method. The multiscale structure of the potential V need not be known in order to apply the method. Thus, in this section, we consider the potential V in the form where we do not impose the structural assumption (2.4) on the potential function $V : \mathbb{R}^d \rightarrow \mathbb{R}$.

The new, penalized mass-matrix of the system is the position dependent tensor defined in (1.3). The associated modified impulses are denoted

$$p_\nu = M_\nu(q)M^{-1}p. \quad (3.1)$$

When ν becomes large, the velocities are bound to remain tangent to the manifolds $\{q|\xi(q) = z\}$, and orthogonal motions are arbitrarily slowed down. Conversely, when $\nu = 0$, one recovers the original highly oscillatory system. Since the modification in M_ν depends on the position q , new geometry is introduced and an additional correction (1.4) in the potential energy is required in order to preserve original statistics in the position variable. This correction is in fact close to the standard Fixman corrector for ν large (see (7.5)). Defining $G(q)$ as the $n \times n$ Gram matrix associated with the fast degrees of freedom

$$G(q) = \nabla_q^T \xi M^{-1} \nabla_q \xi, \quad (3.2)$$

one has the following property of the correcting potential.

PROPOSITION 3.1. *Up to an additive constant, we have*

$$V_{\text{fix},\nu}(q) = \frac{1}{2\beta} \ln \det \left(G(q) + \frac{1}{\nu^2} M_z^{-1} \right), \quad (3.3)$$

and thus (up to additive constants)

$$\lim_{\nu \rightarrow +\infty} V_{\text{fix},\nu} = V_{\text{fix}} = \frac{1}{2\beta} \ln \det (G(q)), \quad \text{and} \quad \lim_{\nu \rightarrow 0} V_{\text{fix},\nu} = 0.$$

Proof. Using the identity for a non-diagonal matrix J of dimension $n_1 \times n_2$:

$$\det (\text{Id}_{n_1} + J J^T) = \det (\text{Id}_{n_2} + J^T J),$$

one observes

$$\det (M_\nu) = \det (M) \det (\nu^2 M_z) \det \left(G + \frac{1}{\nu^2} M_z^{-1} \right)$$

from which the expression for the corrected Fixman potential follows. \square

The associated modified Hamiltonian is then given by

$$H_\nu(p_\nu, q) = \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu + V(q) + V_{\text{fix},\nu}(q), \quad (3.4)$$

and $H_0 = H$ is the original Hamiltonian (1.1).

By construction, statistics of positions q of the mass penalized Hamiltonian are independent of the penalization, leading to the *exact canonical statistics* in position variables.

PROPOSITION 3.2 (Exact statistics). *The canonical distribution associated with the mass-penalized Hamiltonian (3.4) is given by*

$$\mu_\nu(dp_\nu dq) = \frac{1}{Z_\nu} e^{-\beta H_\nu(p_\nu, q)} dp_\nu dq. \quad (3.5)$$

Its marginal probability distribution in q is

$$\frac{1}{Z_\nu} \int e^{-\beta H_\nu(p_\nu, q)} dp_\nu = \frac{e^{-\beta V(q)} dq}{\int e^{-\beta V(q)} dq}$$

which is the original canonical distribution (2.2) in the position variables, and is independent of the mass penalization parameter ν .

Proof. The normalization of Gaussian integrals in the p_ν variables yields

$$\int e^{-\beta \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu} dp_\nu = \left(\frac{2\pi}{\beta} \right)^{d/2} \sqrt{\det(M_\nu)},$$

which is cancelled out by the Fixman corrector $V_{\text{fix}, \nu}$ and the result follows. \square

Sampling such a system can be done using the standard Langevin stochastic perturbation as detailed in Definition 2.1. However, the direct discretization of the equation of motion given by H_ν (e.g., by an explicit scheme) is bound to be unstable (from non-linear instabilities) when the fast degrees of freedom are not affine functions. In order to construct stable schemes one may rather use an implicit formulation of the Hamiltonian (3.4), in conjunction with a solver which enforces the constraints. To obtain such a formulation we extend the state space with n new variables (z_1, \dots, z_n) , and associated moments $(p_{z_1}, \dots, p_{z_n})$. The auxiliary mass-matrix for the new degrees of freedom is then given by M_z . The new extended Hamiltonian of the system H_{IMMP} , defined by (1.5), is now defined in $\mathbb{R}^{d+n} \times \mathbb{R}^{d+n}$, where n position constraints denoted by (C_ν) are included. This construction implies n hidden constraints on momenta. The equivalence of the two Hamiltonians (3.4) and (1.5) formulations is stated as a simple separate lemma.

LEMMA 3.3. *The equations of motion associated with the penalized mass-matrix Hamiltonian (3.4) or the extended Hamiltonian with constraints (1.5) are identical.*

Proof. The Lagrangian associated with H_{IMMP} is given by

$$L_{\text{IMMP}}(\dot{q}, \dot{z}, q, z) = \frac{1}{2} \dot{q}^T M \dot{q} + \frac{1}{2} \dot{z}^T M_z \dot{z} - V(q) - V_{\text{fix}, \nu}(q),$$

and includes hidden constraints on velocities $\dot{z} = \nu \nabla_q^T \xi \dot{q}$ implied by the constraints (C_ν) on position variables. Replacing \dot{z} and z in L_{IMMP} by their expressions as functions of \dot{q} and q , one obtains the Lagrangian associated with H_ν . \square

The stochastically perturbed equations of motion of the Langevin type associated with (1.5) define the dynamics with implicit mass-matrix penalization.

DEFINITION 3.4 (IMMP). *The implicit Langevin process associated with Hamiltonian H_{IMMP} and constraints (C_ν) is defined by the following equations of motion*

$$\begin{cases} \dot{q} = M^{-1}p \\ \dot{z} = M_z^{-1}p_z \\ \dot{p} = -\nabla_q V(q) - \nabla_q V_{\text{fix}, \nu}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \dot{p}_z = -\gamma_z \dot{z} + \sigma_z \dot{W}_z + \frac{\dot{\lambda}}{\nu} \\ \xi(q) = \frac{z}{\nu}, \end{cases} \quad (C_\nu) \quad (3.6)$$

The process \dot{W} (resp. \dot{W}_z) is a standard multi-dimensional white noise, γ (resp. γ_z) a $d \times d$ (resp. $n \times n$) non-negative symmetric dissipation matrix, σ (resp. σ_z) is the fluctuation matrix satisfying $\sigma\sigma^T = \frac{2}{\beta}\gamma$ (resp. $\sigma_z\sigma_z^T = \frac{2}{\beta}\gamma_z$). The processes $\lambda \in \mathbb{R}^n$ are Lagrange multipliers associated with the constraints (C_ν) and adapted with the white noise.

This process is naturally equivalent to the explicit mass-penalized Langevin process in $\mathbb{R}^d \times \mathbb{R}^d$ associated with M_ν . Moreover, when the penalization vanishes ($\nu \rightarrow 0$), the evolution law of the process $\{p_t, q_t\}_{t \geq 0}$ or $\{(p_\nu)_t, q_t\}_{t \geq 0}$ converges towards the original dynamics.

PROPOSITION 3.5. *The stochastic process with constraints (3.6) is well-posed and equivalent to the Langevin diffusion in $\mathbb{R}^d \times \mathbb{R}^d$ (see Definition 2.1), with the mass-penalized Hamiltonian H_ν (3.4), and the dissipation matrix given by*

$$\gamma_\nu(q) = \gamma + \nu^2 \nabla_q \xi \gamma_z \nabla_q^T \xi.$$

Furthermore, the process is reversible and ergodic with respect to the canonical distribution (3.5) (with marginal in position variables given by the original potential, i.e., up to the normalization, $e^{-\beta V(q)} dq$).

Proof. Imposing the constraints implies $\nabla_q^T \xi M^{-1} p = \frac{1}{\nu} M_z^{-1} p_z$. Thus by the definition of p_ν we have

$$p_\nu = p + \nu \nabla_q \xi p_z.$$

Since the position process $\{q_t\}_{t \geq 0}$ is of finite variation, a short computation shows that for each coordinate $i = 1, \dots, d$

$$\dot{p}_\nu^i = \dot{p}^i + \nu (\partial_{q_i})^T \xi \dot{p}_z + \nu^2 \dot{q}^T \nabla_q (\partial_{q_i} \xi) p_z. \quad (3.7)$$

Furthermore,

$$-\partial_{q_i} \left(\frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu \right) = \partial_{q_i} \left(\frac{1}{2} \dot{q}^T M_\nu \dot{q} \right) = \nu^2 \dot{q}^T \nabla_q (\partial_{q_i} \xi) M_z \nabla_q^T \xi \dot{q},$$

and thus

$$\dot{p}_\nu = \dot{p} + \nu \nabla_q \xi \dot{p}_z - \nabla_q \left(\frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu \right).$$

Substituting the expressions for \dot{p} and \dot{p}_z from (3.6) into (3.7) we obtain

$$\dot{p}_\nu = -\frac{1}{2} p_\nu^T \nabla_q M_\nu^{-1} p_\nu - \nabla_q V(q) - \nabla_q V_{\text{fix}, \nu}(q) - \gamma \dot{q} - \nu \nabla_q \xi \gamma_z \dot{z} + \sigma \dot{W} + \nu \nabla_q \xi \sigma_z \dot{W}_z, \quad (3.8)$$

which yields the result. \square

PROPOSITION 3.6 (Small penalty). *When $\nu \rightarrow 0$ the evolution law of the processes $\{p_t, q_t\}_{t \geq 0}$ or $\{(p_\nu)_t, q_t\}_{t \geq 0}$ defined by the implicit equations (3.6) converges (in the sense of probability distributions on continuous paths endowed with the uniform convergence) towards the process solving the original Langevin dynamics (2.3).*

Proof. The stochastic differential equation defined by $\dot{q} = M_\nu^{-1} p_\nu$ and (3.8) has smooth coefficients which depend on ν in a continuous fashion ($\nu \mapsto M_\nu$ and $\nu \mapsto V_{\text{fix}, \nu}$ are continuous). Standard results on weak convergence ([16]) of stochastic processes imply the result as stated. \square

When the mass penalty tends to infinity, the IMMP process converges to a constrained process on the manifold $\mathcal{M}_{z_{t=0}} = \{q \mid \xi(q) = z_{t=0}\}$ (surface measures are described in Section A).

PROPOSITION 3.7 (Large penalty). *Consider a family of initial conditions indexed by ν and satisfying*

$$\sup_\nu |\nu (\xi(q_{t=0}) - z_{t=0})| < +\infty,$$

and assume that the Gram matrix G is invertible in a neighborhood of $\mathcal{M}_{z_{t=0}}$. Then when $\nu \rightarrow +\infty$ the IMMP Langevin stochastic process (3.6) converges weakly towards the decoupled limiting processes with constraints

$$\begin{cases} \dot{q} = M^{-1}p, \\ \dot{p} = -\nabla_q V - \nabla_q V_{\text{fix}} - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ \xi(q) = z_{t=0}, \\ \dot{z} = M_z^{-1}p_z, \\ \dot{p}_z = -\gamma_z \dot{z} + \sigma_z \dot{W}_z. \end{cases} \quad (C) \quad (3.9)$$

where $\{\lambda_t\}_{t \geq 0}$ are adapted stochastic processes defining the Lagrange multipliers associated with the constraints (C).

Furthermore, the process $\{q_t, p_t\}_{t \geq 0}$ defines an effective dynamics with constraints (see also Definition 7.2) on the sub-manifold $\mathcal{M}_{z_{t=0}}$. It is reversible with respect to its stationary canonical distribution given, up to the normalization, by

$$e^{-\beta(H(p,q)+V_{\text{fix}}(q))} \sigma_{T^*\mathcal{M}_{z_{t=0}}}(dp dq)$$

with the q -marginal $e^{-\beta V(q)} \delta_{\xi(q)=0}(dq)$. When γ and γ_z are strictly positive definite the process is ergodic.

Proof. By a simple translation, it is sufficient to show the proposition for $z_{t=0} = 0$. Satisfying the constraint (C_ν) in (3.6) implies a hidden constraint in the momentum space, $\nabla_q \xi M^{-1}p = \frac{1}{\nu} M_z^{-1}p_z$. Differentiating this expression with respect to time and replacing the result in (3.6) yields an explicit formula for the Lagrange multipliers

$$\dot{\lambda} = (G + \frac{1}{\nu^2} M_z^{-1})^{-1} \left[\text{Hess}(\xi)(M^{-1}p, M^{-1}p) + \nabla_q \xi M^{-1}f_q - \frac{1}{\nu} M_z^{-1}f_z \right], \quad (3.10)$$

with forces (f_q, f_z)

$$\begin{aligned} f_q &= -\nabla_q V - \nabla_q V_{\text{fix},\nu} - \gamma M^{-1}p + \sigma \dot{W}, \\ f_z &= -\gamma_z M_z^{-1}p_z + \sigma_z \dot{W}_z, \end{aligned}$$

and the Hessian $\text{Hess}(\xi)$ of the mapping ξ acting on the velocities $M^{-1}p$. This calculation shows that (3.6) is in fact a standard stochastic differential equation with smooth coefficients, and thus has a unique strong solution. The coefficients of these stochastic differential equations are continuous in the limit $\frac{1}{\nu} \rightarrow 0$, at least in a δ -neighborhood of \mathcal{M}_0 in which G is invertible. The formally computed limiting process is given by (3.9) with the Lagrange multipliers solving

$$\dot{\lambda} = G^{-1} (\text{Hess}(\xi)(M^{-1}p, M^{-1}p) + \nabla_q^T \xi M^{-1}f_q).$$

By construction, this limiting process satisfies the constraint $\xi(q) = 0$. Its coefficients are Lipschitz and the process is well-posed. As a result of those properties, the rigorous proof of weak convergence follows classical arguments, see [16], that are divided into three steps

- (i) We truncate the process (3.6) to a compact neighborhood of \mathcal{M}_0 .
- (ii) The continuity of the Markov generator with respect to $\frac{1}{\nu}$ implies tightness for the associated $\frac{1}{\nu}$ -sequences of truncated processes with the limit being uniquely defined by (3.9).
- (iii) The limiting process remains on \mathcal{M}_0 , which implies weak convergence of the sequence without truncation.

The process (3.9) is thus a Langevin process with constraints, exhibiting reversibility properties with respect to the associated Boltzmann canonical measure and is ergodic when γ is strictly positive definite (see the

summary in Appendix B). Note that the q -marginal is geometrically corrected by the Fixman potential term. \square

We conclude this section by discussing some consequences for numerical computations.

REMARK 3.8.

- (i) *Proposition 3.6 and 3.7 imply that IMMP scheme is a tunable interpolation between the exact dynamics (2.3), and the dynamics with constraints (3.9).*
- (ii) *In Section 7 where a stiffness parameter, corresponding to a large penalty ν , is introduced for slow/fast systems leads to a Markovian effective motion constrained on the slow manifold \mathcal{M}_0 , usually seen as a heuristically consistent Markovian approximation of the exact dynamics (see [33]).*

4. Numerical integration. The key ingredient for achieving efficient numerical simulation is to use an integrator that enforces the constraints associated with the implicit formulation of the mass penalized dynamics (3.6). The implicit structure of (3.6) leads to numerical schemes that are potentially asymptotically stable in stiff cases (Section 7). On the other hand, when the penalization ν vanishes with the time-step the scheme becomes consistent with respect to the original exact dynamics (2.3). One may then consider the mass-penalization introduced here as a special method of pre-conditioning for a stiff ODE system with an “implicit”, in the time evolution sense, structure. Here, the “implicit” structure amounts to solving the imposed constraints $\xi(q) = \frac{z}{\nu}$ in (3.6).

It lies outside the scope of this paper to make a review of standard numerical methods for constrained mechanical systems, we refer to [14] as a classical textbook, and to the series ([38, 25, 15, 21, 8]) as a sample of works on practical developments of numerical methods. The IMMP method is presented with the classical leapfrog/Verlet scheme that enforces constraints, usually called RATTLE in (4.2). It can be implemented by a simple modification of standard schemes constraining fDOFs. The scheme is second order, reversible and symplectic. This choice is largely a presentation matter, for practical purposes one can refer to one’s favorite numerical integrator for Hamiltonian systems with or without stochastic perturbations.

As an option, one can also add a Metropolis acceptance/rejection corrector at each time step. If the underlying integrator is reversible and preserves the phase-space measure, this extension leads to a scheme which *exactly* preserves canonical distributions. Such algorithms are usually referred to as Hybrid Monte Carlo (HMC). Generically HMC are sampling algorithms which use the underlying dynamics of the system to generate moves in the configuration space, which are accepted or rejected according to a Metropolis rule. While this method may allow for larger time steps in the integrator it may also lead to higher computational costs when the acceptance ratio becomes low. The acceptance ratio often decreases when the dimension of the system increases ([24]). However, many improvements and modifications ([11, 22, 2, 23]) of the HMC algorithm have been developed since its introduction in [12, 13]. While the HMC methods were initially used in simulations applied to quantum statistical field theories, they have been also employed to wide range of simulations in macromolecular systems, e.g., [9, 36].

We implement numerical discretization of the Langevin process with constraints (3.6) obtained by splitting the Hamiltonian part and the Gaussian fluctuation/dissipation perturbation. The idea of using a Hybrid Monte Carlo step in discretizing a Langevin process goes back to [22]. Note that by using a Metropolis acceptance/rejection rule (HMC), or simply by weighting statistical averages, the forces associated with the Fixman corrector need not be computed when one is only interested in sampling.

We recall that we consider the IMMP dynamics (3.6), which consists of a Langevin process with constraints defined by the following elements

1. the Hamiltonian H_{IMMP} defined in (1.5),
2. the dissipation matrix $\begin{pmatrix} \gamma & 0 \\ 0 & \gamma_z \end{pmatrix}$,
3. the inverse temperature β .

REMARK 4.1. *A useful variant, when using HMC strategies, consists in modifying the Hamiltonian*

H_{IMMP} (1.5) in the integrator (importance sampling) by neglecting the Fixman corrector

$$\begin{cases} \tilde{H}_{\text{IMMP}}(p, p_z, q, z) = \frac{1}{2}p^T M^{-1}p + \frac{1}{2}p_z^T M_z^{-1}p_z + \tilde{V}(q) \\ \xi(q) = \frac{z}{\nu} \end{cases} \quad (C_\nu), \quad (4.1)$$

where \tilde{V} is a potential that may be chosen arbitrarily (typically $\tilde{V} = V$). Indeed, only the underlying phase space structure, which does not depend on \tilde{V} , is necessary in HMC methods. The correct potential $V + V_{\text{fix},\nu}$ has only to be used in the Metropolis step, or by weighting ensemble averages, to ensure exact canonical sampling. Thus potentially costly evaluations of the gradient of the Fixman corrector $V_{\text{fix},\nu}$ can be avoided.

SCHEME 4.1 (Leapfrog/Verlet algorithm for Langevin IMMP (3.6)).

Step 1: Integrate the Hamiltonian part with:

$$\begin{cases} p_{n+1/2} = p_n - \frac{\delta t}{2}(\nabla_q V + \nabla_q V_{\text{fix},\nu})(q_n) - \nabla_q \xi(q_n)\lambda_{n+1/2} \\ p_{n+1/2}^z = p_n^z + \frac{1}{\nu}\lambda_{n+1/2} \\ q_{n+1} = q_n + \delta t M^{-1}p_{n+1/2} \\ z_{n+1} = z_n + \delta t M_z^{-1}p_{n+1/2}^z \\ \xi(q_{n+1}) = \frac{z_{n+1}}{\nu} \end{cases} \quad (C_{1/2}) \quad (4.2)$$

$$\begin{cases} p_{n+1} = p_{n+1/2} - \frac{\delta t}{2}(\nabla_q V + \nabla_q V_{\text{fix},\nu})(q_{n+1}) - \nabla_q \xi(q_{n+1})\lambda_{n+1} \\ p_{n+1}^z = p_{n+1/2}^z + \frac{1}{\nu}\lambda_{n+1} \\ \nabla_q^T \xi(q_{n+1}) M^{-1}p_{n+1} = \frac{1}{\nu}M_z^{-1}p_{n+1}^z \end{cases} \quad (C_1).$$

Step 2: Integrate the Gaussian fluctuation/dissipation part with a mid-point Euler scheme with constraints (see Appendix C).

To obtain exact sampling by correcting the time step errors, one needs to introduce a Metropolis correction.

SCHEME 4.2 (HMC).

Step 1: Compute $(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z)$ with (4.2), and set $\Delta H_{n+1} = H_{\text{IMMP}}(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z) - H_{\text{IMMP}}(q_n, z_n, p_n, p_n^z)$.

Step 2: Accept the step with the probability $\min(1, e^{-\beta \Delta H_{n+1}})$, otherwise reverse impulses and set

$$(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z) = (q_n, z_n, -p_n, -p_n^z).$$

Step 3: Integrate the Gaussian fluctuation/dissipation part with a mid-point Euler scheme (for details see Appendix C).

This numerically constructed Markov chain preserves the canonical distribution.

PROPOSITION 4.2 (Exact sampling). Assume that (4.2) is defined globally. The numerical discretization of (3.6) described in Scheme 4.2 generates a Markov chain that leaves the canonical distribution (3.5) invariant, with marginal distribution in position variables is the original distribution $e^{-\beta V(q)}dq$, which is independent of the mass-penalization ν .

Proof. The statement follows from reversibility and measure preserving properties of Verlet schemes (see [14]), from the Hybrid Monte Carlo rule ([13]), and from the construction of the mass-penalized Hamiltonian (Proposition 3.5). \square

One can now construct consistent schemes by letting the penalty $\nu = \bar{\nu}\delta t^k$ go to zero with the time-step, for some $k > 0$. Indeed, Proposition 3.6 shows that the mass-penalized dynamics (3.6) converges towards the exact original dynamics for $\nu = 0$ at order ν^2 . Consequently most of the usual numerical schemes will be consistent at their own approximation order, but bounded above by $2k$. Neglecting the order of the

fluctuation/dissipation part (Step 2 of Scheme 4.1), we deduce the following convergence property. The order of convergence refers to the maximal integer k such that the convergence of trajectories with respect to the uniform norm occurs at the rate $\mathcal{O}(\delta t^k)$.

PROPOSITION 4.3 (Time-step consistency). *Assume that the numerical flow (4.2) is defined globally, and that $\nu = \bar{\nu}\delta t^k$. Then the IMMP numerical scheme (4.2) is consistent of the order $\min(2, 2k)$ with respect to the original exact limiting process (2.3) without thermostat ($\gamma = 0$, $\sigma = 0$).*

Proof. We consider the scheme (4.2) in the variables $(p, q, \nu p_z, \nu z)$. Then the force field only depends on q and the global mass-matrix is smooth with respect to ν^2 . The sub-manifold defining the constraints $(\nu^2 \xi(q) = \nu z)$ is also smooth with respect to ν^2 . By the implicit function theorem we have that locally the RATTLE scheme is the standard leapfrog scheme (see [14, 34]), and the local mapping depends smoothly on ν^2 . Therefore the standard calculation of the order of the leapfrog scheme holds uniformly with respect to ν , [14]. Now, the mass-penalized dynamics (3.6) is a differential equation (deterministic here) whose coefficients differ from the original process by a smooth perturbation of order ν^2 . The result thus follows from applying a simple Gronwall argument. \square

Note also that taking $\nu \rightarrow +\infty$ in (4.2) gives a numerical scheme with constraints consistent with the limiting constrained dynamics of Proposition 3.7. For slow/fast systems (see Section 7.2), large penalty will lead to asymptotic stability in the stiffness limit.

We conclude this section with a practical recipe for tuning the mass matrix penalty. This can be done for instance by computing the time averaged Metropolis acceptance ratio $r = \mathbb{E}[e^{-\beta(\Delta H_n)}]$ in Scheme 4.2, which gives a precise quantification of the time-step error. Then increasing the penalty ν can save computational time as long as it leads to an increase of the average ratio r . Indeed, this means that the selected fDOFS are limiting the time-step stability region. Prescribing a time average ratio r (for instance 0.90), a maximal time-step δt_{max} associated with the largest penalty ν_{max} that is able to improve stability can be obtained in this way. Finally, one can set, for example, $\nu = \frac{\nu_{max}}{\delta t_{max}}\delta t$ in Scheme 4.2 to obtain an order 2 convergent scheme with increased stability region.

5. A high dimensional numerical test case. The behavior of numerical methods for Hamiltonian systems becomes of particular interest when simulating high-dimensional systems. Such systems in general exhibit many different time-scales and one seeks to go beyond stability constraints implied by the fastest microscopic degrees of freedom. As mentioned in the introduction, globally implicit methods are usually too costly, while efficient splitting methods using minimal implicitness represent an active research area. In order to demonstrate applicability of the IMMP method we consider non-convex repulsive interactions between particles in the gas phase which makes implicit methods intractable.

On the other hand, direct constraints on microscopic fast degrees of freedom are commonly used in practice and can introduce a potentially strong bias in the macroscopic behavior of the whole system. In the case presented here, a particle chain with pairwise interactions, one may be interested in some macroscopic observable, like the total length of the chain. Direct constraints would then lead to a *totally rigid chain*, losing completely the evolution of the latter macroscopic variable.

We numerically study the IMMP strategy described in Section 3 and Section 4, and the latter is systematically compared with the exact dynamics numerically integrated by a simple Leapfrog/Verlet scheme with a thermostat. Langevin thermostats are used, with a Metropolis/HMC step for the dynamics integrator as described by Scheme 4.2. An appropriate scaling of the chain of size N at the mass transport level with a large mass-penalty of order $\nu := \bar{\nu}N$ are chosen. Numerical results show that IMMP induces a gain in time stepping of order N , while the macroscopic dynamics (in particular convergence to equilibrium), given by a formal non-linear stochastic partial differential equation, remains of order 1. Macroscopic dynamics of the IMMP method also stays close to the exact Verlet integration when re-scaled penalty $\bar{\nu}$ is small.

5.1. The particle chain model. The model we consider consists of a chain of particles which interact through a non-convex repulsive pairwise potential of the form

$$\begin{cases} v_{\text{int}}(r) = f(r), & r \leq r_{\text{int}}, \\ v_{\text{int}}(r) = 0, & r > r_{\text{int}}, \end{cases} \quad (5.1)$$

where $f(r_{\text{int}}) = f'(r_{\text{int}}) = 0$ and $\lim_{r \rightarrow -\infty} f(r) = +\infty$. Each particle is also individually submitted to a macroscopic confining exterior potential. After converting to the non-dimensional form the typical quantities involved in the model enable us to write a scaling at the mass-transport level where the dynamics of the chain is described by the Hamiltonian

$$H_N(q, p) = \frac{1}{2} p^T p + \sum_{i=1}^{N-1} v_{\text{int}}(\nabla_i^d q) + \sum_{i=1}^N v_{\text{ext}}(q_i), \quad (5.2)$$

and by a coupling with an exterior thermal bath at the re-scaled inverse temperature

$$\beta_N := \beta N^{-1}. \quad (5.3)$$

In the expression (5.2) the functions $r \in \mathbb{R} \mapsto v_{\text{int}}(r) \in \mathbb{R}$ and $q \mapsto v_{\text{ext}}(q) \in \mathbb{R}$ are the smooth interaction potential and the exterior potential respectively. The linear operator $\nabla^d : \mathbb{R}^N \rightarrow \mathbb{R}^{N-1}$ having the components

$$\nabla_i^d q = \frac{q_{i+1} - q_i}{1/N}, \quad i = 1, \dots, N-1,$$

represents the discrete gradient associated to the chain with the Neumann boundary conditions. Its transpose operator is denoted $(\nabla^d)^T : \mathbb{R}^{N-1} \rightarrow \mathbb{R}^N$. The particles are represented by their re-scaled positions $q = (q_1, \dots, q_N)$, so that the typical position and deviation of q is formally of order 1 with respect to N . This can be seen by considering particles in the chain as indexed by $x = \frac{i}{N} \in [0, 1]$. The thermodynamic limit $N \rightarrow \infty$ of the Hamiltonian (5.2) then converges to

$$\lim_{N \rightarrow \infty} \frac{1}{N} H_N(q, p) = \int_0^1 \frac{1}{2} p(x)^2 dx + \int_0^1 v_{\text{int}}\left(\frac{dq}{dx}\right) dx + \int_0^1 v_{\text{ext}}(q(x)) dx, \quad (5.4)$$

when $x \mapsto q(x)$ is a smooth function on $[0, 1]$ satisfying the Neumann boundary conditions $q'(0) = q'(1) = 0$. We choose to work with the scaling given by (5.2)-(5.4) in order to prescribe the macroscopic timescale of the chain profile at order 1 with respect to N . Using the matrix notation, we write the stochastically perturbed equations of motion

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -(\nabla^d)^T v'_{\text{int}}(\nabla^d q) - v'_{\text{ext}}(q) - \gamma p + \sqrt{N} \sigma \dot{W} \end{aligned}$$

with the fluctuation/dissipation identity $\sigma \sigma^T = 2\beta^{-1}\gamma$.

Next we turn to the mass-penalized Hamiltonian model for this particle chain. The number of independent fast degrees of freedom that will be penalized is equal to $N-1$. The penalized degrees of freedom are given by the inter-particle distances $\xi_i(q) = z_i = q_i - q_{i-1}$. The mass-penalization intensity has to scale formally as

$$\nu := \nu_N = \bar{\nu} N$$

to give a discrete operator meaning to the perturbed mass-matrix in \mathbb{R}^N

$$M_{\nu_N} = \text{Id} + \bar{\nu}^2 (\nabla^d)^T M_z \nabla^d.$$

When $M_z = \text{Id}$ the matrix perturbation is given by the discrete Laplacian

$$\Delta_d := -(\nabla^d)^T \nabla^d.$$

with the Neumann boundary conditions. Following our general construction we obtain the mass-penalized Hamiltonian

$$H_{\nu_N}(p_{\nu_N}, q) = \frac{1}{2} p_{\nu_N}^T M_{\nu_N}^{-1} p_{\nu_N} + \sum_{i=1}^N v_{\text{int}}(\nabla_i^d q) + \sum_{i=1}^N v_{\text{ext}}(q_i), \quad (5.5)$$

and the stochastically perturbed equations of motion with $M_z = \text{Id}$ are

$$\ddot{q} = (\text{Id} - \bar{\nu}^2 \Delta_d)^{-1} \left(-(\nabla^d)^T v'_{\text{int}}(\nabla^d q) - v'_{\text{ext}}(q) - \gamma \dot{q} + \sigma \sqrt{N} \dot{W} \right). \quad (5.6)$$

The form (5.6) of the equations of motion clearly demonstrates the regularizing effect of the mass-matrix penalization. Indeed, all the high frequencies generated by the microscopic forces $-(\nabla^d)^T v'_{\text{int}}(\nabla^d q)$ are *filtered* by a “low-pass filter” represented by the regularizing operator $(\text{Id} - \bar{\nu}^2 \Delta_d)^{-1}$.

5.2. Numerical results. In this section we demonstrate behavior of the IMMP method with the Hamiltonian (5.5) for different values of the re-scaled mass-matrix penalization intensity $\bar{\nu}$, the time step parameter δt , and the system size N . In particular, we perform a systematic comparison with the exact dynamics (5.2) integrated by the Leapfrog/Verlet scheme.

The simulations were carried out with the re-scaled quadratic exterior potential

$$v_{\text{ext}}(q) = \left(\frac{q - 0.5}{2.2} \right)^2,$$

and with a double-well repulsive re-scaled interaction potential

$$\begin{cases} v_{\text{int}}(r) = (50((r - 0.1)^2 - 0.05^2))^2, & r \leq 0.1 \\ v_{\text{int}}(r) = 0, & r > 0.1, \end{cases}$$

and the rescaled inverse temperature $\beta_N = 10 N^{-1}$. We choose the penalizing mass-matrix to be the identity matrix $M_z = \text{Id}$, and the dissipation parameter $\gamma = 0.1$.

TEST I: *Comparison of macroscopic dynamics.*

We demonstrate approximation of the macroscopic dynamics by the IMMP method simulating two observables: (i) the length of the chain: $l = q_N - q_1$, and (ii) the center of mass: $c = q_{N/2}$. The dynamical behavior for short times of the chain length l and the center of mass c for different values of the mass-matrix penalization intensity $\bar{\nu}$ is depicted in Figure 5.1. The same realization of the noise is used in each simulation, and no Metropolis rejection is used. One can observe the smoothing induced by the IMMP method, however, typical evolution timescales are not changed by the IMMP method.

Next, the convergence to equilibrium of l and c are analyzed. The initial condition is taken out of equilibrium with zero potential energy and particles concentrated near the central position $q = 0.5$. The equilibrium probability density functions (PDF) of l and c are computed for the IMMP case and the exact dynamics in Figures 5.2, exact sampling ensuring identity of the two calculations. The PDFs are estimated using a Gaussian kernel estimator. The evolution of the relative entropy between a trajectory sample and the equilibrium distribution is plotted in Figure 5.3. We observe that the convergence of the IMMP dynamics is slightly faster than the exact dynamics with the Verlet integrator. Convergence to equilibrium on shorter times can be also observed by inspecting the auto-correlations in time series of $l(t)$ and $c(t)$ in Figure 5.3. Convergence of the IMMP method is again faster. However, it is interesting to note that the

mass penalization introduces an inertial artifact that produces oscillations in the auto-correlations of the center of mass evolution.

The average transition time $\tau(\bar{\nu}, N)$ of the chain center of mass c between $q = 0.4$ and $q = 0.6$ is summarized in Table 5.1¹ for different values of the chain size N and the mass penalization $\bar{\nu}$. The results show that the average time $\tau(\bar{\nu}, N)$ of the macroscopic dynamics of c scales as $\mathcal{O}(1)$ with respect to the system size, and is continuous with respect to the exact dynamics when $\bar{\nu} \rightarrow 0$.

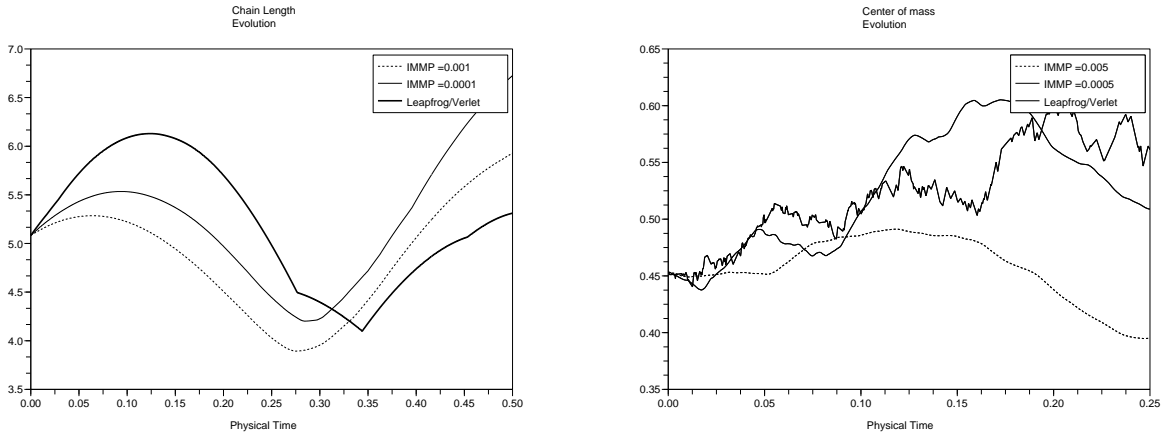


FIGURE 5.1. Short time evolution of the chain length and the center of mass for different values of the IMMP mass-penalization $\bar{\nu}^2 = \{10^{-3}, 10^{-4}\}$ and $\bar{\nu}^2 = \{5 \cdot 10^{-3}, 5 \cdot 10^{-4}\}$, $N = 500$.

	$\bar{\nu}^2 = 10^{-2}$	$\bar{\nu}^2 = 10^{-3}$	$\bar{\nu}^2 = 10^{-4}$	$\bar{\nu}^2 = 10^{-5}$	Leapfrog/Verlet
$N = 100$	1.6 (1)	1.3 (1)	1.2 (1)	1.0 (1)	1.0 (1)
$N = 300$	1.5 (1)	1.3 (1)	1.2 (1)	1.1 (1)	1.0 (1)
$N = 500$	1.4 (1)	1.2 (1)	1.2 (1)	1.1 (1)	1.0 (1)

TABLE 5.1

Average transition time of the chain center of mass from the position $q_{N/2} = 0.4$ to the position $q_{N/2} = 0.6$, simulated by the IMMP method.

TEST II: Relaxation of time-step restrictions.

We first observe in Figure 5.5(a) the relationship between the time-step δt and the mass penalty $\bar{\nu}$. The average Metropolis acceptance rate is computed for different time-steps and different values of the mass-matrix penalization. The critical time-step, for which the average acceptance rate is approximately 0.5, is significantly increased by the mass-matrix penalization, and is broadly proportional to the re-scaled mass-matrix penalty $\bar{\nu}$.

In Figure 5.5(b), the critical time-step is studied for different values of the chain size N and of the mass-matrix penalization, and it is compared to the exact dynamics. Using linear regression the dependence of the critical time-step on the system size is estimated and compared to the exact Verlet integration in Table 5.2¹. These results are in a good agreement with Proposition 6.3 that characterizes the dependence for the linear case.

Conclusions.

¹Digits inside the brackets indicate the typical error of the last digit.

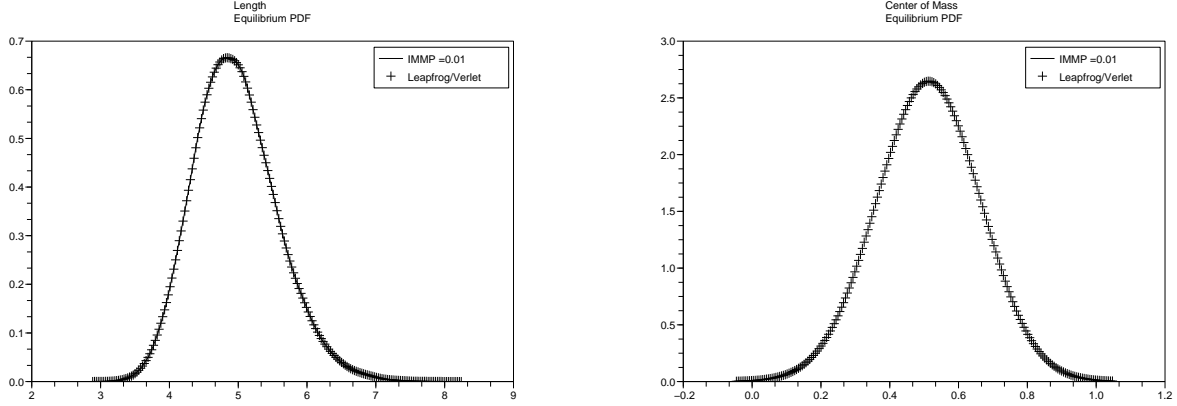


FIGURE 5.2. *Equilibrium PDF of the chain length and the center of mass. Simulated by the IMMP method with the penalty $\bar{\nu}^2 = 10^{-2}$.*

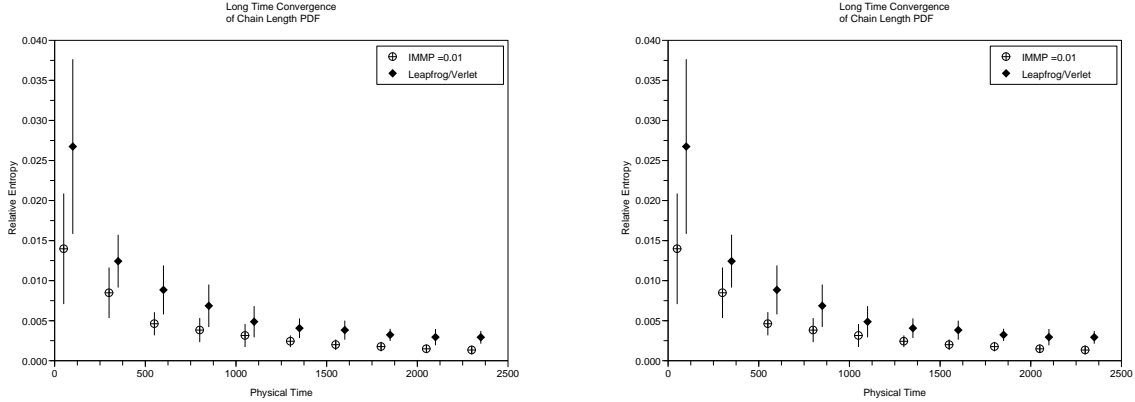


FIGURE 5.3. *Convergence of the relative entropy of the PDF of a trajectory of the chain length and the center of mass towards equilibrium. Simulated by the IMMP method with the penalty $\bar{\nu}^2 = 10^{-2}$. The error bar represents extreme quartiles as computed over 200 realizations.*

The presented numerical studies demonstrate that for a prescribed macroscopic timescale of the chain, the IMMP allows for arbitrary large (of order N) relaxation of the time-step restriction. This observation shows that for a large system, $N \rightarrow \infty$, the IMMP method can achieve arbitrarily large computational savings when compared to the standard Verlet algorithm used for integrating the exact dynamics.

6. Numerical analysis of the particle chain. In this section, the particle chain model (5.2) is analyzed in the special case of harmonic interactions. We consider the thermodynamic limit $N \rightarrow +\infty$ where N is the size of the system. It is shown that the macroscopic dynamics of the IMMP method behaves continuously (uniformly with N , and in the L_2 norm for the position profile) with respect to the re-scaled mass penalty parameter $\bar{\nu}$.

At the same time, the time-step stability of the IMMP numerical scheme (4.1) is compared with the standard Verlet scheme, and the critical time step is shown to be increased by a factor $\bar{\nu}N$.

From the spectral point of view, the IMMP method behaves in this linear case as a low-pass filter. This proves, in this simplified case, the ability of IMMP method to respect macroscopic dynamical equivalence,

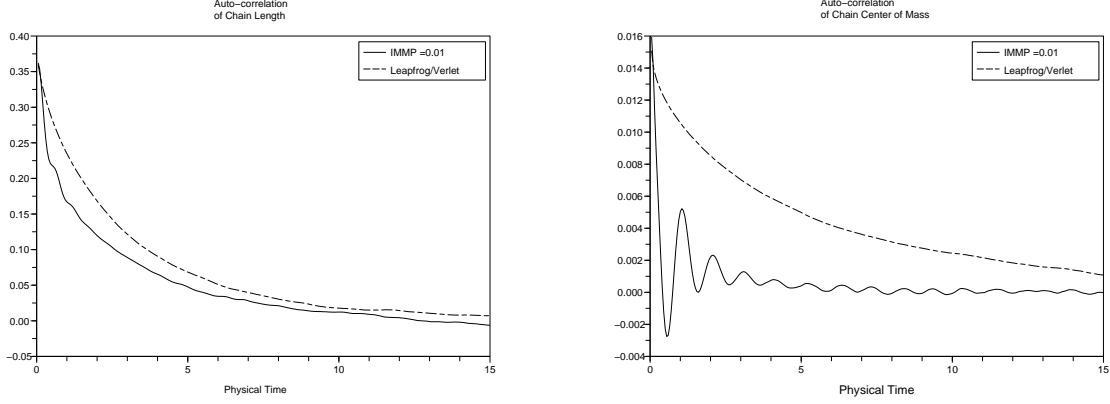


FIGURE 5.4. Auto-correlation of a trajectory of the chain length and the center of mass. Simulated by the IMMP method with the penalty $\bar{\nu}^2 = 10^{-2}$.

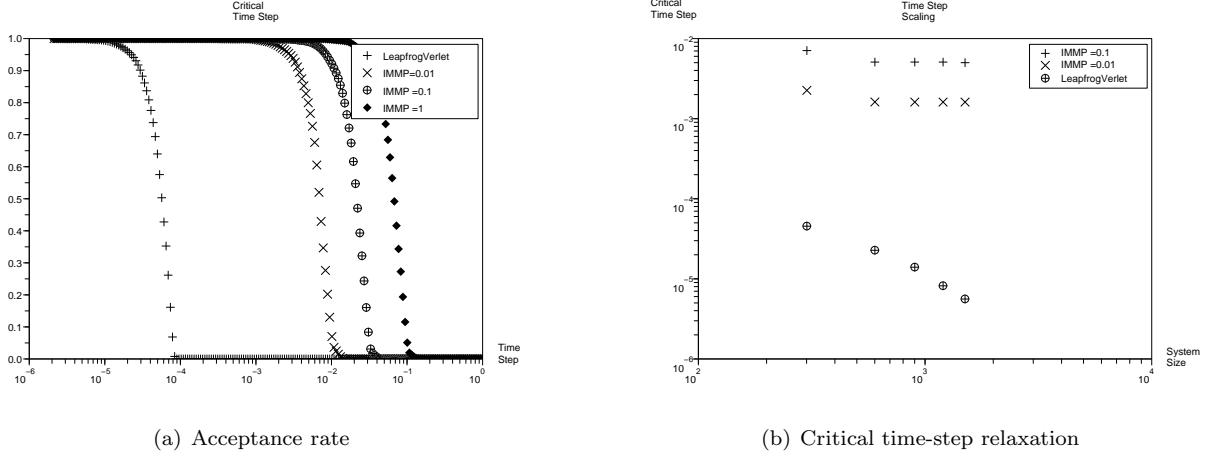


FIGURE 5.5. (a) The average acceptance rate of the Metropolis step for different values of the mass-matrix penalization $\bar{\nu}^2 = \{1, 10^{-1}, 10^{-2}\}$ and the system size $N = 100$. (b) Critical time-step that achieves the Metropolis acceptance rate of 0.5. Comparison between the IMMP dynamics ($\bar{\nu}^2 = \{10^{-1}, 10^{-2}\}$), and the exact dynamics, for different values of the system size N .

while saving computational time up to a factor of order $\mathcal{O}(N)$.

6.1. Conservation of macroscopic dynamics. We consider the penalized Hamiltonian (5.5) when the interaction potential is quadratic $v_{\text{int}}(r) = r^2/2$. The penalizing matrix is the identity matrix $M_z = \text{Id}$, hence the penalized mass-tensor becomes $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$, and the fluctuation/dissipation tensor is taken proportional to the identity matrix. The system of equations (5.6) then becomes

$$\begin{cases} \dot{q} = (\text{Id} - \bar{\nu}^2 \Delta_d)^{-1} p \\ \dot{p} = \Delta_d q - v'_{\text{ext}}(q) - \gamma \dot{q} + \sigma \sqrt{N} \dot{W}, \end{cases} \quad (6.1)$$

with fluctuation/dissipation identity $\gamma = 2\sigma^2\beta^{-1}$. The associated canonical equilibrium distribution is then given by the re-scaled inverse temperature $\beta_N = \beta N^{-1}$.

	$\bar{\nu}^2 = 10^{-1}$	$\bar{\nu}^2 = 10^{-2}$	Leapfrog/Verlet
$\delta t \propto N^{-\alpha}$	$\alpha = 0.2(1)$	$\alpha = 0.2(1)$	$\alpha = 1.2(1)$

TABLE 5.2

Scaling of the critical time-step for the Metropolis acceptance rate $r = 0.5$.

In order to treat the limit $N \rightarrow \infty$, we introduce the ℓ_2 -norm in the position space

$$\|q\|_{\ell_2}^2 := \frac{1}{N} \sum_{i=1}^N q_i^2 = \frac{1}{N} q^T q,$$

as well as the h_{-1} -norm in the momentum space

$$\|p\|_{h_{-1}}^2 = \left\| (-\Delta_d)^{-1/2} \left(p - \frac{1}{N} \sum_{i=1}^N p_i \right) \right\|_{\ell_2}^2 + \left(\frac{1}{N} \sum_{i=1}^N p_i \right)^2.$$

In the above expression, $\frac{1}{N} \sum_{i=1}^N p_i$ can be seen as the orthogonal projection in ℓ^2 on the one dimensional kernel of the Neumann discrete Laplacian Δ_d . $\|q\|_0^2 + \|p\|_{-1}^2$ is a quadratic form that endows the phase-space with a Hilbert space structure.

PROPOSITION 6.1 (Convergence of the macroscopic dynamics). *Assume that the exterior potential v_{ext} is bounded and that its derivative satisfies the Lipschitz condition*

$$\|v'_{\text{ext}}(q_2) - v'_{\text{ext}}(q_1)\|_{h_{-1}} \leq L_v \|q_2 - q_1\|_{\ell_2},$$

where L_v is independent of N . For any $T > 0$, let $t \mapsto (p^\bar{\nu}(t), q^\bar{\nu}(t))$ be the solution for $t \in [0, T]$ of the evolution equation (6.1) with the initial condition

$$(p^\bar{\nu}(0), q^\bar{\nu}(0)) = (M_{\nu_N}^{-1/2} p^0(0), q^0(0)),$$

where $(p^0(0), q^0(0))$ is distributed according to the original equilibrium canonical distribution (associated with (5.2)-(5.3)). Then for all $t \in [0, T]$ one has the uniform convergence

$$\lim_{\bar{\nu} \rightarrow 0} \limsup_{N \rightarrow +\infty} \mathbb{E} \left[\|q^\bar{\nu}(t) - q^{\bar{\nu}=0}(t)\|_{\ell_2}^2 \right] = 0.$$

Proof. We write $X = (q, p)$, and introduce the norm:

$$\|X\|_{\bar{\nu}} = \|q\|_{\ell_2} + \left\| M_{\nu_N}^{-1/2} p \right\|_{h_{-1}}.$$

The system (6.1) becomes a stochastic differential equation in the form

$$dX_t^\bar{\nu} = A_{\bar{\nu}} X_t^\bar{\nu} + F(X_t^\bar{\nu}) + \Sigma dW_t, \quad (6.2)$$

where by definition

$$A_{\bar{\nu}} := \begin{pmatrix} 0 & (\text{Id} - \bar{\nu}^2 \Delta_d)^{-1} \\ \Delta_d & 0 \end{pmatrix}, \quad F(X) := \begin{pmatrix} 0 \\ -v'_{\text{ext}}(q) - \gamma p \end{pmatrix}, \quad \Sigma := \begin{pmatrix} 0 \\ \sqrt{N} \sigma \end{pmatrix}.$$

Duhamel formula gives an implicit expression for differences of solutions of (6.2) with the same noise

$$\begin{aligned} X_t^\bar{\nu} - X_t^0 &= (e^{A_{\bar{\nu}} t} - e^{A_0 t}) X_0^0 + \int_0^t (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) (F(X_s^0) ds + \Sigma dW_s) \\ &\quad + e^{t A_{\bar{\nu}}} (X_0^\bar{\nu} - X_0^0) + \int_0^t e^{A_{\bar{\nu}}(t-s)} (F(X_s^\bar{\nu}) - F(X_s^0)) ds. \end{aligned} \quad (6.3)$$

We estimate the individual terms on the right hand side in (6.3). We define P as the coordinate transformation associated with the orthonormal spectral decomposition

$$-\Delta_d = P^{-1} \text{diag}(\delta_0, \dots, \delta_{N-1}) P,$$

where $PP^T = \text{Id}$, and the eigenvalues of the discrete Neumann Laplacian are given, for $k = 0, \dots, N-1$, by

$$\delta_k = 4N^2 \sin^2\left(\frac{k\pi}{2N}\right) \underset{N \rightarrow \infty}{\sim} k^2 \pi^2. \quad (6.4)$$

Denoting the spectral coordinates

$$\hat{X} = (\hat{q}, \hat{p}) = (N^{-1/2} P p, N^{-1/2} P q)$$

we have

$$\|X\|_{\bar{\nu}}^2 = \hat{p}_0^2 + \sum_{k=1}^{N-1} \frac{\delta_k}{1 + \bar{\nu}^2 \delta_k} \hat{p}_k^2 + \sum_{k=0}^{N-1} \hat{q}_k^2.$$

The spectral decomposition leads to a block diagonal form of the operator $e^{A_{\bar{\nu}} t}$ with diagonal 2×2 blocks in the spectral basis

$$e^{\widehat{A}_{\bar{\nu}}^{(0)} t} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

as well as for $k = 1, \dots, N-1$

$$e^{\widehat{A}_{\bar{\nu}}^{(k)} t} = \begin{pmatrix} 0 & (1 + \bar{\nu}^2 \delta_k)^{-1} \\ \delta_k & 0 \end{pmatrix},$$

where $\widehat{A}_{\bar{\nu}}^{(k)}$ is the 2×2 block associated with the coordinates (\hat{q}_k, \hat{p}_k) . Since $\widehat{A}_{\bar{\nu}}^{(k)}$ conserves the k -mode energy $\delta_k(\hat{q}_k)^2 + (1 + \bar{\nu}^2 \delta_k)^{-1}(\hat{p}_k)^2$, one can check that for any $N \geq 1$

$$\|e^{A_{\bar{\nu}} t}\|_{\bar{\nu}}^2 \leq 2 + 2t^2.$$

Similarly, since in the sense of symmetric matrices $M_{\nu_N}^{-1/2} \leq \text{Id}$, we have the bound

$$\|F(X^{\bar{\nu}}) - F(X^0)\|_{\bar{\nu}} \leq \left\| M_{\nu_N}^{-1/2} (v'_{\text{ext}}(q^{\bar{\nu}}) - v'_{\text{ext}}(q^0) + \gamma(p^{\bar{\nu}} - p^0)) \right\|_{h_{-1}} \leq (L_F + \gamma) \|X^{\bar{\nu}} - X^0\|_{\bar{\nu}}.$$

Using random independence of Brownian increments we compute

$$\mathbb{E} \left[\left\| \int_0^t (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) \Sigma dW_s \right\|_{\bar{\nu}}^2 \right] = \int_0^t \sum_{i=1}^N \left\| (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) \Sigma_{\cdot, i} \right\|_{\bar{\nu}}^2 ds.$$

Applying Gronwall lemma in (6.3) and collecting all terms we obtain

$$\mathbb{E} \left[\|X_t^{\bar{\nu}} - X_t^0\|_{\bar{\nu}}^2 \right] \leq C_T \left(\mathbb{E} \left[\|X_0^{\bar{\nu}} - X_0^0\|_{\bar{\nu}}^2 + m_T \right] \right), \quad (6.5)$$

where C_T is independent of N , and with X^0 being distributed canonically m_T is given

$$m_T = \sup_{t \in [0, T]} \left(\mathbb{E} \left[\|(e^{A_{\bar{\nu}} t} - e^{A_0 t}) X^0\|_{\bar{\nu}}^2 \right] + \mathbb{E} \left[\|(e^{A_{\bar{\nu}} t} - e^{A_0 t}) F(X^0)\|_{\bar{\nu}}^2 \right] + \sum_{i=1}^N \|(e^{A_{\bar{\nu}} t} - e^{A_0 t}) \Sigma_{\cdot, i}\|_{\bar{\nu}}^2 \right).$$

For a given random vector X such that $\mathbb{E} [\|X\|_0^2] < +\infty$, Parseval identity and the inequality $\|\cdot\|_{\bar{\nu}} \leq \|\cdot\|_0$ imply

$$\mathbb{E} \left[\left\| (e^{\widehat{A}_{\bar{\nu}} t} - e^{\widehat{A}_0 t}) X \right\|_{\bar{\nu}}^2 \right] = \sum_{k=1}^{N-1} \mathbb{E} \left[\left\| (e^{\widehat{A}_{\bar{\nu}} t} - e^{\widehat{A}_0 t}) \hat{X} \right\|_{k, \bar{\nu}}^2 \right] \leq 2 \sum_{k=1}^{N-1} \mathbb{E} \left[\left\| \hat{X} \right\|_{k, 0}^2 \right], \quad (6.6)$$

where $\|\cdot\|_{k, \bar{\nu}}$ is the restriction to the k -th mode (\hat{q}_k, \hat{p}_k) . Then one has, by orthogonality of P ,

$$\sum_{i=1}^N \mathbb{E} \left[\left\| \widehat{\Sigma_{\cdot, i}} \right\|_{k, 0}^2 \right] = \sigma^2 \sum_{i=1}^N P_{k, i}^2 \frac{1}{\delta_k} \leq \frac{\sigma^2}{\delta_k}.$$

Now the canonical distribution of $X^0 = (q^0, p^0)$ can be described as follows. Up to normalization, the distribution of X^0 has the density $e^{-\frac{\beta}{N} \sum_{i=1}^d v_{\text{ext}}(q_i^0)}$ with respect to the Gaussian distribution with the covariance matrix $\beta^{-1} \text{Id}$ for momenta variables, and the covariance matrix $(\beta \Delta_d)^{-1}$ for positions. Thus we have the bound

$$\mathbb{E} \left[\left\| \widehat{X^0} \right\|_{k, 0}^2 \right] \leq 2e^{4\beta \|v_{\text{ext}}\|_{\infty}} \frac{1}{\delta_k \beta},$$

as well as

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\left\| \widehat{F(X^0)} \right\|_{k, 0}^2 \right] \leq e^{4\beta \|v_{\text{ext}}\|_{\infty}} \mathbb{E} \left[\left\| \mathcal{F} \circ F \circ \mathcal{F}^{-1}(\widehat{G^0}) \right\|_{k, 0}^2 \right],$$

where \mathcal{F} denotes the Fourier series expansion on $[0, 1]$ with Neumann conditions, and $(\widehat{G^0}_k)_{k \geq 1}$ are canonical centered Gaussian i.i.d. variables with the covariance matrix $\beta^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{k^2 \pi^2} \end{pmatrix}$. By the Lipschitz assumption the series is bounded

$$\sum_{k=1}^{+\infty} \mathbb{E} \left[\left\| \mathcal{F} \circ F \circ \mathcal{F}^{-1}(\widehat{G^0}) \right\|_{k, 0}^2 \right] \leq (L_v + \gamma) \mathbb{E} \left[\left\| \widehat{G^0} \right\|_0^2 \right] = \sum_{k=1}^{+\infty} \frac{2(L_v + \gamma)}{\beta k^2 \pi^2}.$$

Since $\lim_{\bar{\nu} \rightarrow 0} \|e^{\widehat{A}_{\bar{\nu}}^{(k)} t} - e^{\widehat{A}_0^{(k)} t}\| = 0$, one can take the limit $N \rightarrow +\infty$ and use the uniform convergence of the series in (6.6) to obtain $\lim_{\bar{\nu} \rightarrow 0} \lim_{N \rightarrow +\infty} m_T = 0$ in (6.5). The convergence of the initial condition $\lim_{\bar{\nu} \rightarrow 0} \lim_{N \rightarrow +\infty} \mathbb{E} \|X_0^{\bar{\nu}} - X_0^0\|_{\bar{\nu}}^2$ follows by using similar arguments. The proof is complete. \square

6.2. Relaxation of time-step stability restriction. To demonstrate improved stability properties of time integration algorithms we consider the IMMP scheme (4.2) associated with the mass-matrix penalized Hamiltonian (5.5). Note that when the constraints are linear, the leapfrog scheme (RATTLE) applied to an implicit Hamiltonian is identical to the usual leapfrog scheme for the associated explicit Hamiltonian (5.5). We restrict the rigorous analysis to the quadratic interaction potential ($v_{\text{int}}(y) = \frac{y^2}{2}$), zero exterior potential ($v_{\text{ext}} = 0$), and to the mass-matrix penalization operator $(\text{Id} - \bar{\nu}^2 \Delta_d)$. The leapfrog scheme is defined as

$$\begin{cases} p_{n+1/2} = p_n + \frac{\delta t}{2} (-\Delta_d) q_n \\ q_{n+1} = q_n + \delta t M_{\nu_N}^{-1} p_{n+1} \\ p_{n+1} = p_{n+1/2} + \frac{\delta t}{2} (-\Delta_d) q_{n+1} . \end{cases}$$

Denoting the spectral variables for $k = 1, \dots, N - 1$

$$\begin{cases} \hat{v}^k = \left(\frac{\delta_k}{1 + \bar{\nu}^2 \delta_k} \right)^{1/2} \sqrt{N} P p \\ \hat{x}^k = \left(\frac{1 + \bar{\nu}^2 \delta_k}{\delta_k} \right)^{1/2} \sqrt{N} P q, \end{cases} \quad (6.7)$$

we write

$$\begin{pmatrix} \hat{v}_{n+1}^k \\ \hat{x}_{n+1}^k \end{pmatrix} = L_k \begin{pmatrix} \hat{v}_n^k \\ \hat{x}_n^k \end{pmatrix},$$

where

$$L_k = \begin{pmatrix} 1 - \frac{h_k^2}{2} & -h_k + \frac{h_k^3}{4} \\ h_k & 1 - \frac{h_k^2}{2} \end{pmatrix}, \quad \text{and} \quad h_k = \delta t \frac{\delta_k^{1/2}}{(1 + \bar{\nu}^2 \delta_k)^{1/2}}.$$

Since $\det(L_k) = 1$, the standard CFL stability condition is equivalent to

$$|\text{Tr}(L_k)| \leq 2$$

which is fulfilled if and only if $h_k \leq 2$ for all $k \leq N - 1$. Thus we arrive at the following bound on the time step

$$\delta t \leq 2 \min_{0 \leq k < N} \left(\frac{1 + \bar{\nu}^2 \delta_k}{\delta_k} \right)^{1/2}.$$

Summarizing the above calculations and recalling (6.4) we have the following characterization of the stability properties.

PROPOSITION 6.2. *Suppose $v_{\text{ext}} = 0$ and consider a harmonic interaction potential $v_{\text{int}}(r) = \frac{r^2}{2}$ with the mass-matrix penalization $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$. The leapfrog/Verlet integration of the IMMP harmonic Hamiltonian (5.2) is stable in the spectral sense if and only if*

$$\delta t \leq \left(4\bar{\nu}^2 + \frac{1}{N^2 \sin^2 \left(\frac{(N-1)\pi}{2N} \right)} \right)^{1/2}. \quad (6.8)$$

Since we work with a Metropolis correction of the hybrid Monte-Carlo type, we are also interested in the limiting behavior of the energy variation compared to the temperature, i.e.,

$$\beta_N (H(p_{n+1}, q_{n+1}) - H(p_n, q_n)),$$

when (p_n, q_n) are distributed according to the canonical distribution. This quantity gives the average acceptance rate of the Metropolis correction. The result we present here is similar to [4] where the authors analyze infinite dimensional sampling with the standard Metropolis-Hastings Markov chains.

PROPOSITION 6.3. *Suppose $v_{\text{ext}} = 0$ and consider a harmonic interaction potential $v_{\text{int}}(r) = \frac{r^2}{2}$ with the mass-matrix penalization $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$. Suppose the state variable $X = (p_{\nu_N}, q)$ is a random variable distributed according to the canonical distribution associated with the mass-matrix penalized Hamiltonian (5.2). Then the energy variation $\beta_N \Delta H$ after one step of the leapfrog integration scheme converges in distribution, up to normalization and centering, to the Gaussian random variable*

$$\frac{\beta_N \Delta H - m_N}{\sigma_N} \xrightarrow[N \rightarrow +\infty]{\text{Law}} \mathcal{N}(0, 1),$$

with the mean and variance in the infinite size asymptotics for the IMMP method $\bar{\nu} > 0$ and $\delta t \equiv \delta t_N = o(1)$

$$m_N \underset{N \rightarrow +\infty}{\sim} \frac{N \delta t_N^6}{32 \bar{\nu}^6}, \quad \text{and} \quad \sigma_N^2 \underset{N \rightarrow +\infty}{\sim} \frac{N \delta t_N^6}{16 \bar{\nu}^6},$$

and for the Verlet integration of exact dynamics with $\delta t \equiv \delta t_N = o(1/N)$

$$m_N \underset{N \rightarrow +\infty}{\sim} \frac{5}{8} N^7 \delta t_N^6, \quad \text{and} \quad \sigma_N^2 \underset{N \rightarrow +\infty}{\sim} \frac{5}{4} N^7 \delta t_N^6.$$

Proof. We start with a canonically distributed state $X = (q, p)$, which is, by assumption on the form of the interaction potential, a Gaussian random vector. After changing to the spectral coordinates (6.7) we have the spectral representation of the Hamiltonian

$$\beta_N H = \beta \sum_{k=1}^{N-1} \frac{\delta_k^{1/2}}{2(1 + \bar{\nu}^2 \delta_k)^{1/2}} ((\hat{v}^k)^2 + (\hat{x}^k)^2),$$

and introducing Gaussian random vectors U and V with the identity covariance matrix we can write

$$\hat{x}^k = \beta^{-1/2} \frac{(1 + \bar{\nu}^2 \delta_k)^{1/4}}{\delta_k^{1/4}} U_k, \quad \text{and} \quad \hat{v}^k = \beta^{-1/2} \frac{(1 + \bar{\nu}^2 \delta_k)^{1/4}}{\delta_k^{1/4}} V_k.$$

We then compute explicitly the change of the Hamiltonian after one step of the leapfrog integration

$$\beta_N \Delta H = \sum_{k=1}^{N-1} \frac{1}{2} \begin{pmatrix} U_k \\ V_k \end{pmatrix}^T (L_k^T L_k - \text{Id}) \begin{pmatrix} U_k \\ V_k \end{pmatrix}. \quad (6.9)$$

Since $\det(L_k^T L_k) = 1$ the matrix $L_k^T L_k - \text{Id}$ has two positive eigenvalues $(\lambda_k - 1, 1/\lambda_k - 1)$ which satisfy

$$\begin{aligned} \lambda_k + 1/\lambda_k - 2 &= \text{Tr}(L_k^T L_k - \text{Id}) = \frac{h_k^6}{16}, \\ (\lambda_k - 1)^2 + (1/\lambda_k - 1)^2 &= \text{Tr}(L_k^T L_k)^2 - 2\text{Tr}(L_k^T L_k) = \frac{h_k^{12}}{256} + \frac{h_k^6}{8}. \end{aligned}$$

Combing with (6.9) we find

$$m_N \equiv \mathbb{E}[\beta_N \Delta H] = \sum_{k=1}^{N-1} \frac{h_k^6}{2^5}, \quad \text{and} \quad \sigma_N^2 \equiv \text{Var}[\beta_N \Delta H] = \sum_{k=1}^{N-1} \frac{h_k^6}{2^4} + \frac{h_k^{12}}{2^9}.$$

Moreover, the Lindenberg or simply Lyapunov condition in the general central limit theorem (see [18]) is verified since we work with a sum of χ^2 random variables, thus concluding the first part of the proof.

Recalling

$$h_k = \delta t \frac{\sin(\frac{k}{N} \frac{\pi}{2})}{(\frac{1}{4N^2} + \bar{\nu}^2 \sin^2(\frac{k}{N} \frac{\pi}{2}))^{1/2}},$$

we compute the convergent Riemann sums for $p = 6$ and $p = 12$. For the case $\bar{\nu} \neq 0$ we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^{N-1} h_k^p = \frac{\delta t^p}{\bar{\nu}^p}.$$

If $\bar{\nu} = 0$ we obtain

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N^{p+1}} \sum_{k=1}^{N-1} h_k^p &= \lim_{N \rightarrow \infty} \frac{\delta t^p}{N} \sum_{k=1}^{N-1} 2^p \sin^p \left(\frac{k}{2N} \pi \right) \\ &= \delta t^p 2^p \int_0^1 \sin^p \left(\frac{\pi}{2} x \right) dx. \end{aligned}$$

Thus for $p = 6$ we have that the series sums to $20\delta t^6$. Then the asymptotic behavior follows from the assumption $\delta t_N^{12} \ll \delta t_N^6$, and similarly $\delta t_N^{12} N^{12} \ll \delta t_N^6 N^6$ in the case $\bar{\nu} = 0$. \square

REMARK 6.4. The two propositions proved in this section characterize the restrictions imposed by the stability of the resulting scheme. In Proposition 6.2, stability in the large system size limit, $N \rightarrow +\infty$, is equivalent to the inequality (6.8). In this case the restriction of the time-step size is imposed by the numerical integrator. On the other hand the stability for the scheme which uses a Metropolis corrector is linked to the acceptance rate of the Metropolis step. In Proposition 6.2, stability in the large system size limit is equivalent to the non-vanishing Metropolis acceptance rate, which is equivalent to bounded from above average energy variation m_N and bounded variance σ_N of the energy variation. In either case, the IMMP method ($\bar{\nu} > 0$) induces a relative increase of order N for the boundary of numerical stability as compared to the exact dynamics $\bar{\nu} = 0$ integrated with the Verlet scheme.

7. The stiff limit. Throughout this section, one introduces a potential function with an explicit dependence with respect to the fast variables $(q, z) \mapsto U(q, z)$ together with a stiffness parameter ϵ . The potential energy V can then be written in the form

$$V(q) = U\left(q, \frac{\xi(q)}{\epsilon}\right),$$

with a confining assumption

$$\inf_{q \in \mathbb{R}^d} U(q, z) \geq K(z),$$

where $\lim_{z \rightarrow \infty} K(z) = +\infty$ strictly faster than at any logarithmic rate. The functions ξ are then indeed “fast” degrees of freedom (fDOFs) in the limit $\epsilon \rightarrow 0$, the system being confined to the slow sub-manifold $\mathcal{M}_0 = \{q \mid \xi(q) = 0\}$.

We will prove, that under appropriate scaling of the mass penalty $\nu_\epsilon = \frac{\bar{\nu}}{\epsilon}$, the IMMP method is asymptotically stable in the stiff limit, converging towards standard effective dynamics on the slow manifold \mathcal{M}_0 .

7.1. Thermostatted stiff systems. The canonical distribution becomes

$$\mu_\epsilon(dp dq) = \frac{1}{Z_\epsilon} e^{-\beta(\frac{1}{2}p^T M^{-1}p + U(q, \frac{\xi(q)}{\epsilon}))} dp dq. \quad (7.1)$$

In the infinite stiffness limit ($\epsilon \rightarrow 0$) the measure concentrates on the slow manifold \mathcal{M}_0 . The limit is computed using the co-area formula (see Appendix A for relevant definitions of surface measures). In order to characterize the limiting measure we introduce the effective potential

$$U_{\text{eff}}(q) = -\frac{1}{\beta} \ln \int e^{-\beta U(q, z)} dz. \quad (7.2)$$

LEMMA 7.1. *In the infinite stiffness limit ($\epsilon \rightarrow 0$), the highly oscillatory canonical distribution (7.1) converges $\mu_\epsilon \rightarrow \mu_0$ (in distribution) towards $\mu_0(dp dq)$, which is supported on \mathcal{M}_0 , and defined as*

$$\mu_0(dp dq) = \frac{1}{Z_0} e^{-\beta(\frac{1}{2}p^T M^{-1}p + V_{\text{eff}}(q))} dp \delta_{\xi(q)=0}(dq). \quad (7.3)$$

Its marginal distribution in position is given, up to the normalization, by

$$e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq). \quad (7.4)$$

Proof. It is sufficient to consider distributions in the position variable q only. Let \mathcal{U}^δ be a δ -neighborhood of \mathcal{M}_0 where

$$dq = \epsilon^n \delta_{\xi(q)=\epsilon z}(dq) dz.$$

We construct a decomposition $\varphi = \varphi_1 + \varphi_2$ of continuous bounded observables such that $\text{supp } \varphi_1 \subset \mathcal{U}^\delta$ and $\text{supp } \varphi_2 \cap \mathcal{U}^{\delta/2} = \emptyset$. Using the confining property of $U(q, \cdot)$ we obtain

$$\int \varphi(q) e^{-\beta U(q, \frac{\xi(q)}{\epsilon})} dq = \epsilon^n \int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq) dz + \mathcal{O}(e^{-\beta K(\delta/2\epsilon)}).$$

By continuity of $\epsilon \mapsto \int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq)$ and by the dominated convergence theorem

$$\int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq) dz \rightarrow \int \varphi_1(q) e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq) = \int \varphi(q) e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq),$$

and the result follows after normalization. \square

The infinite stiffness limit ($\epsilon \rightarrow 0$) of highly oscillatory dynamics has been studied in a series of papers [35, 37, 27, 5, 32, 33]. The limiting dynamics can be fully characterized in special cases. For example, when the highly oscillatory potential is linear and non-resonant (at least almost everywhere on the trajectory, see [37]), it can be described through adiabatic effective potentials. See also [10, 6] for a recent work on some related numerical issues. However, when the system is thermostatted, one can postulate an “ad hoc” effective dynamics ([33]) exhibiting the appropriate limiting canonical distribution given by (7.4). Such dynamics can be obtained by constraining the system to the slow manifold \mathcal{M}_0 , and adding a correcting entropic potential, sometimes called “Fixman” corrector from [19], which is due to the geometry of \mathcal{M}_0 , and is given by

$$V_{\text{fix}}(q) = \frac{1}{2\beta} \ln(\det G(q)), \quad (7.5)$$

where $G(q)$ is the $n \times n$ Gram matrix defined in (3.2).

In general, since the effective potential (7.2) is not explicit, one may need to couple the system with virtual fast degrees of freedom to enforce the appropriate effective dynamics associated with (7.2). The resulting extended Hamiltonian is then defined on the state space $T^*(\mathcal{M}_0 \times \mathbb{R}^n)$ (the cotangent bundle) and is given by

$$\begin{cases} H_{\text{eff}}(p, p_z, q, z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + U(q, z) + V_{\text{fix}}(q) \\ \xi(q) = 0. \end{cases} \quad (C) \quad (7.6)$$

DEFINITION 7.2 (Effective Langevin process with constraints). *The constrained Langevin process associated with Hamiltonian (7.6) is defined by the following stochastic differential equations*

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{z} = M_z^{-1} p_z \\ \dot{p} = -\nabla_1 U(q, z) - \nabla_q V_{\text{fix}}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_2 \xi \dot{\lambda} \\ \dot{p}_z = -\nabla_z U(q, z) - \gamma_z \dot{z} + \sigma_z \dot{W}_z \\ \xi(q) = 0, \end{cases} \quad (C) \quad (7.7)$$

where ∇_1 and ∇_2 are respectively derivatives with respect to the first and second variable of the function $U(q, z)$, \dot{W} (resp. \dot{W}_z) is the standard multi-dimensional white noise, γ (resp. γ_z) a $d \times d$ (resp. $n \times n$) symmetric positive semi-definite dissipation matrix, σ (resp. σ_z) is the fluctuation matrix satisfying $\sigma\sigma^T = \frac{2}{\beta}\gamma$ (resp. $\sigma_z\sigma_z^T = \frac{2}{\beta}\gamma_z$). The processes $\lambda \in \mathbb{R}^n$ are Lagrange multipliers associated with the constraints (C) and adapted with respect to the white noise.

We formulate reversibility of this process as a separate lemma.

LEMMA 7.3. *The process defined in (7.7) is reversible with respect to the associated canonical distribution whose marginal distribution in (q, p) variables is*

$$\mu_{\text{eff}}(dp dq) = \frac{1}{Z_{\text{eff}}} e^{-\beta(\frac{1}{2}p^T M^{-1}p + V_{\text{eff}}(q) + V_{\text{fix}}(q))} \sigma_{T^*\mathcal{M}_0}(dp dq) \quad (7.8)$$

with the q -marginal

$$e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq).$$

When γ and γ_z are strictly positive definite, the process is ergodic.

Proof. The process (7.7) is a Langevin process with mechanical constraints, exhibiting reversibility properties with respect to the associated Boltzmann canonical measure (see the summary in Appendix B). Then the q -marginal is obtained by remarking that the integration of any function of $\frac{1}{2}p^T M^{-1}p + \frac{1}{2}p_z^T M_z^{-1}p_z$ with respect to $dp_z \sigma_{T_q^*\mathcal{M}_0}(dp)$ results in a constant independent of q . \square

The properties of thermostatted highly oscillatory systems are summarized in Table 7.1.

	Finite stiffness $\epsilon > 0$	Infinite stiffness limit $\epsilon \rightarrow 0$	Infinite stiffness $\epsilon = 0$
Dynamics	Highly oscillatory + fluct./diss.	Adiabatic (if non-resonant) + non-Markov fluct./diss.	Effective with constraints + fluct./diss.
Statistics	Canonical	Positions on \mathcal{M}_0 , free velocities.	Canonical on $T^*\mathcal{M}_0$, geometric corrector.
Numerics	Leapfrog/Verlet + fluct./diss.	Time-step restrictions ($\delta t = o(\epsilon)$)	Leapfrog/Verlet with constraints + fluct./diss.

TABLE 7.1

Stiff Hamiltonian systems and associated commonly used numerical methods (\mathcal{M}_0 denotes the slow manifold). Two different schemes are required for the stiff system and its effective Markovian approximation.

7.2. Stability of the IMMP dynamics. We assume that the mass-matrix penalty parameter $\nu \equiv \nu_\epsilon$ grows to infinity in such a way that

$$\lim_{\epsilon \rightarrow 0} \epsilon \nu_\epsilon = \bar{\nu}.$$

The original Hamiltonian with the stiffness parameter is expressed explicitly as

$$H_\epsilon(p, q) = \frac{1}{2}p^T M^{-1}p + U(q, \frac{\xi(q)}{\epsilon}), \quad (7.9)$$

and including the mass-matrix penalization one gets

$$H_{\nu_\epsilon}(p_{\nu_\epsilon}, q) = \frac{1}{2}p_{\nu_\epsilon}^T M_{\nu_\epsilon}^{-1}p_{\nu_\epsilon} + U(q, \frac{\xi(q)}{\epsilon}) + V_{\text{fix}, \nu_\epsilon}(q), \quad (7.10)$$

or in its implicit formulation

$$\begin{cases} H_{\text{IMMP}}(q, z, p, p_z) = \frac{1}{2}p^T M^{-1}p + \frac{1}{2}p_z^T M_z^{-1}p_z + U(q, \frac{z}{\nu_\epsilon \epsilon}) + V_{\text{fix}, \nu_\epsilon}(q), \\ \xi(q) = \frac{1}{\nu_\epsilon}z. \end{cases} \quad (C_{\nu_\epsilon}) \quad (7.11)$$

One immediately sees that H_{IMMP} is non-singular when $\epsilon \rightarrow 0$ and converges to the effective Hamiltonian on the slow manifold,

$$\begin{cases} H_{\text{eff}, \bar{\nu}}(q, z, p, p_z) = \frac{1}{2}p^T M^{-1}p + \frac{1}{2}p_z^T M_z^{-1}p_z + U(q, \frac{z}{\bar{\nu}}) + V_{\text{fix}}(q) \\ \xi(q) = 0. \end{cases} \quad (C). \quad (7.12)$$

The expression (7.11) represents a minor generalization of H_{eff} (7.6), but it leads to the same canonical marginal distribution $\mu_{\text{eff}}(dp dq)$ in (p, q) variables as given by (7.8). The continuity in ϵ of H_{IMMP} implies stability of the associated dynamics and their numerical integrators. We first derive the limits of the original and penalized canonical distribution.

PROPOSITION 7.4 (Limits of canonical distributions). *Consider the canonical distributions $\mu_{\nu_\epsilon}(dp dq)$ associated with the mass penalized Hamiltonian (7.10), but considered with respect to the $(p = M M_{\nu_\epsilon}^{-1} p_{\nu_\epsilon}, q)$ variables. In the sense of weak convergence of measures we have $\mu_{\nu_\epsilon} \rightarrow \mu_{\text{eff}}$ as $\epsilon \rightarrow 0$ with μ_{eff} defined by (7.8).*

Proof. The first convergence is proved in Lemma 7.1. For the second one, the following notation will be used

$$\begin{cases} \delta_{q, \epsilon z}(dq) = \delta_{\xi(q) = \epsilon z}(dq) \\ \delta_{p, \epsilon p_z}(dp) = \delta_{p^T M^{-1} \nabla \xi(q) = \epsilon M_z^{-1} p_z}(dp). \end{cases}$$

To prove the convergence towards μ_{eff} we consider a δ -neighborhood \mathcal{U}^δ of \mathcal{M}_0 where

$$dp dq = \frac{\epsilon^{2n}}{\det M_z} \delta_{q, \epsilon z}(dq) dz \delta_{p, \epsilon p_z}(dp) dp_z,$$

and a decomposition of the bounded observable (in (p, q) variables) $\varphi = \varphi_1 + \varphi_2$ such that $\text{supp } \varphi_1 \subset \mathcal{U}^\delta$ and $\text{supp } \varphi_2 \cap \mathcal{U}^{\delta/2} = \emptyset$. Thus, keeping in mind that $p_{\nu_\epsilon} = M_{\nu_\epsilon} M^{-1} p$, and using the confining property of the potential $U(q, \cdot)$ we obtain

$$\int \varphi(p, q) e^{-\beta H_{\nu_\epsilon}} dp_{\nu_\epsilon} dq = \int \varphi_1(p, q) e^{-\beta H_{\nu_\epsilon}} dp_{\nu_\epsilon} dq + \mathcal{O}(e^{-\beta K(\delta/\epsilon)}) \equiv I_\epsilon + \mathcal{O}(e^{-\beta K(\delta/2\epsilon)}). \quad (7.13)$$

Applying the change of variables $p_{\nu_\epsilon} = M_{\nu_\epsilon} M^{-1} p$ yields

$$dp_{\nu_\epsilon} = \det(M_{\nu_\epsilon} M^{-1}) dp = \nu_\epsilon^{2n} \det M_z \det(G + \frac{1}{\nu_\epsilon^2} M_z^{-1}) dp,$$

and setting $\epsilon M_z^{-1} p_z = \nabla_q \xi M^{-1} p$ and $\epsilon z = \xi(q)$ we get

$$H_{\nu_\epsilon}(p_{\nu_\epsilon}, q) = \frac{1}{2}p^T M^{-1}p + \nu_\epsilon^2 \epsilon^2 p_z^T M_z^{-1}p_z + U(q, \frac{z}{\nu_\epsilon \epsilon}) + V_{\text{fix}, \nu_\epsilon}(q) = H_{\text{IMMP}}(q, z, p, p_z).$$

Thus substituting back to (7.13) we obtain

$$I_\epsilon = (\nu_\epsilon \epsilon)^{2n} \int \varphi_1 e^{-\beta H_{\text{IMMP}}(q, z, p, p_z)} \det(G + \frac{1}{\nu_\epsilon^2} M_z^{-1}) \delta_{p, \epsilon p_z}(dp) dp_z \delta_{q, \epsilon z}(dq) dz,$$

and thus

$$I_\epsilon \xrightarrow{\epsilon \rightarrow 0} \bar{\nu}^{2n} \int \varphi_1 e^{-\beta H_{\text{eff}, \bar{\nu}}(q, z, p, p_z)} \det(G) \delta_{p, \epsilon p_z}(dp) dp_z \delta_{\xi(q)=0}(dq) dz.$$

Using the co-area formula we obtain

$$\det(G) \delta_{\nabla \xi(q) M^{-1} p=0}(dp) \delta_{\xi(q)=0}(dq) = \sigma_{T^* \mathcal{M}_0}(dp dq),$$

which leads to the final result after integration of the (p_z, z) variables and normalization. \square

REMARK 7.5. Due to the fast oscillations, the distribution of impulses in the limiting distribution μ_0 in (7.3) is uncorrelated, whereas after the mass-matrix penalization, the limiting distribution (3.5) has almost surely co-tangent impulses (i.e., satisfying the constraints $\nabla_q \xi M^{-1} p = 0$). This explains the role of the corrected potential energy V_{fix} taking into account the curvature of \mathcal{M}_0 .

In the next step we inspect the infinite stiffness asymptotic of the penalized dynamics.

PROPOSITION 7.6 (Infinite stiffness limit). *When $\epsilon \rightarrow 0$ with $\nu \equiv \nu_\epsilon \sim \frac{\bar{\nu}}{\epsilon}$ and $V(q, \xi(q)) = U(q, \frac{\xi(q)}{\epsilon})$, the IMMP Langevin stochastic process (3.6) converges weakly towards the following coupled limiting processes with constraints*

$$\begin{cases} \dot{q} = M^{-1} p, \\ \dot{p} = -\nabla_1 U(q, \frac{z}{\bar{\nu}}) - \nabla_q V_{\text{fix}}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ \xi(q) = 0, \\ \dot{z} = M_z^{-1} p_z, \\ \dot{p}_z = -\frac{1}{\bar{\nu}} \nabla_2 U(q, \frac{z}{\bar{\nu}}) - \gamma_z \dot{z} + \sigma_z \dot{W}_z. \end{cases} \quad (C) \quad (7.14)$$

where ∇_1 and ∇_2 are respectively derivatives with respect to the first and second variable of the function $U(q, z)$, and $\{\lambda_t\}_{t \geq 0}$ are adapted stochastic processes defining the Lagrange multipliers associated with the constraints (C).

The process $\{q_t, p_t\}_{t \geq 0}$ defines an effective dynamics with constraints (Definition 7.2) for thermostatted highly oscillatory systems. It is reversible with respect to its stationary canonical distribution given by μ_{eff} (7.8), and is ergodic when (γ, γ_z) are strictly positive definite.

Proof. The proof is similar to the proof of Proposition 3.7. Here we have

$$\nabla_q U = \nabla_1 U + \frac{1}{\epsilon} \nabla_q \xi \nabla_2^T U,$$

and (3.6) translates, up to a change of Lagrange multipliers, into

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{z} = M_z^{-1} p_z \\ \dot{p} = -\nabla_1 U - \nabla_q V_{\text{fix}, \nu}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \dot{p}_z = -\frac{1}{\nu_\epsilon \epsilon} \nabla_2 U - \gamma_z \dot{z} + \sigma_z \dot{W}_z + \frac{\dot{\lambda}}{\nu_\epsilon} \\ \xi(q) = \frac{z}{\nu_\epsilon}, \end{cases} \quad (C_{\nu_\epsilon}). \quad (7.15)$$

The rest follows the proof of Proposition 3.7.

□

REMARK 7.7. When $\bar{\nu} \rightarrow +\infty$, by a classical averaging argument (see, e.g., [27]), one can check that the limiting dynamics are the effective dynamics pointed out in [33]

$$\begin{cases} \dot{q} = M^{-1}p \\ \dot{p} = -\nabla_q U_{\text{eff}}(q) - \nabla_q V_{\text{fix}} - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \xi(q) = 0. \end{cases} \quad (C) \quad (7.16)$$

with the stationary canonical distribution (7.8).

7.3. Stability of the IMMP integrator. The numerical scheme (Scheme 4.1 proposed for the IMMP method (3.6)) is also stable in the limit of infinite stiffness $\epsilon \rightarrow 0$. Recall that we consider a reversible, measure preserving numerical flow $\Phi_{\delta t}^{\nu_\epsilon}(p, p_z, q, z)$ associated with Hamiltonian (7.11) H_{IMMP} with constraints (modified potentials could similarly be considered).

PROPOSITION 7.8 (Asymptotic stability). *In the limit $\epsilon \nu_\epsilon \rightarrow \bar{\nu}$, the numerical flow $\Phi_{\delta t}^{\nu_\epsilon}$ associated with the leapfrog/Verlet integrator with constraints for the IMMP Hamiltonian (7.11) converges towards the numerical flow $\Phi_{\delta t}^{\bar{\nu}}$, which is the leapfrog/Verlet integrator with geometric constraints associated with effective Hamiltonian (7.12) on the slow manifold.*

Proof. The statement is a direct consequence of the implicit function theorem and the continuity of the leapfrog integrator with constraints (4.2) with respect to the parameter $\bar{\nu} = \epsilon \nu_\epsilon$. Indeed, considering the shift of Lagrange multipliers

$$\lambda \rightarrow \lambda + \frac{1}{\epsilon} \nabla_2 U$$

and taking the limit $\epsilon \rightarrow 0$ we obtain the appropriate leapfrog scheme

$$\begin{cases} p_{n+1/2} = p_n - \frac{\delta t}{2} \nabla_1 U(q_n, \frac{z_n}{\bar{\nu}}) - \nabla_q \xi(q_n) \lambda_{n+1/2} \\ p_{n+1/2}^z = p_n^z - \frac{\delta t}{2\bar{\nu}} \nabla_2 U(q_n, \frac{z_n}{\bar{\nu}}) \\ q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2} \\ z_{n+1} = z_n + \delta t M_z^{-1} p_{n+1/2}^z \\ \xi(q_{n+1}) = 0 \\ p_{n+1} = p_{n+1/2} - \frac{\delta t}{2} \nabla_1 U(q_{n+1}, \frac{z_{n+1}}{\bar{\nu}}) - \nabla_q \xi(q_{n+1}) \lambda_{n+1} \\ p_{n+1}^z = p_{n+1/2}^z - \frac{\delta t}{2\bar{\nu}} \nabla_2 U(q_{n+1}, \frac{z_{n+1}}{\bar{\nu}}) \\ \nabla_q \xi(q_{n+1}) M^{-1} p_{n+1} = 0. \end{cases} \quad (C_{1/2}) \quad (C_1)$$

□

By convergence of the Hamiltonian (7.11) to (7.12), similar asymptotic stability properties holds when a Metropolis step is introduced.

The results and properties discussed in this section are summarized in Table 7.2.

	Zero mass penalization $\nu = 0$	Positive mass-penalization $\epsilon, \nu > 0$	Infinite stiffness limit $\epsilon \rightarrow 0, \frac{\nu}{\epsilon} \rightarrow \bar{\nu}$
Dynamics	Highly oscillatory + fluct./diss.	IMMP + fluct./diss.	Effective with constraints+ fluct./diss.
Statistics	Canonical	Canonical with correlated velocities	Canonical on $T^*\mathcal{M}_0$
Numerics	IMMP + fluct./diss.		

TABLE 7.2

The IMMP dynamics and the Verlet numerical integration are both asymptotically stable in the infinite stiffness regime if $\frac{\nu}{\epsilon} \rightarrow \bar{\nu} < +\infty$. If the mass-penalization vanishes ($\nu = 0$) one recovers the original physical stiff system. The canonical distribution is always exact in the position variable. Notice that due to the penalized mass-matrix ($\nu > 0$) the statistics have correlated velocities.

Appendix A. Surface measures. Let \mathbb{R}^d be endowed with the scalar product given by the positive definite matrix M , and consider \mathcal{M}_z a family of sub-manifolds of co-dimension n implicitly defined by n independent functions $\mathcal{M}_z = \{x \in \mathcal{M} \mid \xi_1(q) = z_1, \dots, \xi_n(q) = z_n\}$ for z in a neighborhood of the origin. For each z in a neighborhood of the origin the *conditional measure* $\delta_{\xi(q)=z}(dq)$ is a measure on \mathcal{M}_z defined in such a way that it satisfies the chain rule for conditional expectations with respect to the Lebesgue measure dq , i.e.,

$$dq = \delta_{\xi(q)=z}(dq) dz. \quad (\text{A.1})$$

The *surface measure* $\sigma_{T^*\mathcal{M}_z}(dp)$ is the Hausdorff measure induced by the metric M^{-1} on the co-tangent space $T_q^*\mathcal{M}_z = \{p \mid \nabla_q^T \xi(q) M^{-1} p = 0\}$; and in the same way, $\sigma_{\mathcal{M}_z}(dq)$ is the Hausdorff measure induced by the metric M on the sub-manifold \mathcal{M}_z . It is important to note that, although this is not explicit in notation, σ is defined with respect to the mass-tensor M of the mechanical system. The Liouville measure $\sigma_{T^*\mathcal{M}_z}(dp dq)$ on the co-tangent bundle $T^*\mathcal{M}_z$ is the volume form induced on

$$T^*\mathcal{M}_z = \{(p, q) \mid \nabla_q^T \xi(q) M^{-1} p = 0, \quad \xi(q) = z\}$$

by the usual symplectic form $dp \wedge dq$. It can be described in terms of surface measures as follows

$$\sigma_{T^*\mathcal{M}_z}(dp dq) = \sigma_{T_q^*\mathcal{M}_z}(dp) \sigma_{\mathcal{M}_z}(dq).$$

Finally, the co-area formula (see [17] for a general reference) defines the relative probability density between $\delta_{\xi(q)=z}(dq)$ and $\sigma_{\mathcal{M}_z}(dq)$.

PROPOSITION A.1 (Co-area formula). *Given the invertible Gram matrix associated with the constraints $\xi(q) = z$ in a neighborhood of $\mathcal{M}_z = \{q \mid \xi(q) = z\}$*

$$G(q) = \nabla_q^T \xi M^{-1} \nabla_q \xi,$$

one has

$$\delta_{\xi(q)=z}(dq) = \frac{1}{\sqrt{\det G(q)}} \sigma_{\mathcal{M}_z}(dq).$$

Appendix B. Langevin processes. Defining the Poisson bracket

$$\{\varphi_1, \varphi_2\} = \nabla_p^T \varphi_1 \nabla_q \varphi_2 - \nabla_p^T \varphi_2 \nabla_q \varphi_1,$$

and the dissipation tensor

$$d(q) = \sigma q,$$

where σ is the fluctuation matrix in Definition 2.1, the Markov generator of the Langevin process in Definition 2.1 is

$$\mathcal{L} = \{\cdot, H\} + \frac{1}{\beta} \{d, \{d^T, \cdot\} e^{-\beta H}\} e^{\beta H}.$$

The generator \mathcal{L} satisfies

$$\int \varphi_1 \mathcal{L}(\varphi_2) e^{-\beta H} dp dq = \int \mathcal{L}^*(\varphi_1) \varphi_2 e^{-\beta H} dp dq,$$

where

$$\mathcal{L}^* = \{\cdot, -H\} + \frac{1}{\beta} \{d, \{d^T, \cdot\} e^{-\beta H}\} e^{\beta H}.$$

The generator \mathcal{L}^* defines a Langevin process with the time-reversed Hamiltonian $(-H)$. Reversibility of the process implies that the canonical measure is stationary. Furthermore, if the initial state of the system is a canonically distributed random variable, the probability distribution of a trajectory after the time-reversal is given by a Langevin process with the generator \mathcal{L}^* . When H has the form $H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q)$, reversal of impulses ($p \rightarrow -p$) leads to time-reversed dynamics, and a process with generator \mathcal{L}^* can be constructed by the following simple steps:

1. Reverse momenta ($p \rightarrow -p$).
2. Draw a random path with generator \mathcal{L} .
3. Reverse again momenta ($p \rightarrow -p$).

When holonomic constraints, for instance, of the form

$$\Xi(p, q) = \zeta \Leftrightarrow \begin{cases} p^T M^{-1} \nabla_q \xi = 0 \\ \xi(q) = z \end{cases}$$

are introduced, it is useful to define the Poisson bracket on the co-tangent bundle $T^* \mathcal{M}_z$

$$\{\varphi_1, \varphi_2\}_{\mathcal{M}_z} = \{\varphi_1, \varphi_2\} - \sum_{a,b} \{\varphi_1, \Xi^a\} \Gamma_{a,b}^{-1} \{\Xi^b, \varphi_2\},$$

where Γ is the symplectic Gram matrix of the full constraints

$$\Gamma^{a,b} = \{\Xi^a, \Xi^b\}.$$

As a basic result of symplectic geometry (see [1]), one recovers the divergence formula with respect to the bracket $\{\cdot, \cdot\}_{\mathcal{M}_z}$ and the Liouville measure $\sigma_{T^* \mathcal{M}_z}(dp dq)$

$$\int \{\cdot, \cdot\}_{\mathcal{M}_z} \sigma_{T^* \mathcal{M}_z}(dp dq) = 0.$$

Given a constrained Langevin process in a stochastic differential equation form

$$\begin{aligned} \dot{q} &= \nabla_p H, \\ \dot{p} &= -\nabla_q H - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \end{aligned}$$

where λ are Lagrange multipliers associated with the constraints $\xi(q) = 0$, adapted with respect to the noise \dot{W} , the process $\{p_t, q_t\}_{t \geq 0}$ obeys hidden velocity constraints and is characterized by the stochastic differential equations

$$\begin{aligned}\dot{q} &= \nabla_p H + \nabla_p \Xi \dot{\Lambda}, \\ \dot{p} &= -\nabla_q H - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \Xi \dot{\Lambda},\end{aligned}$$

where Λ are Lagrange multipliers associated with the full constraints $\Xi(p, q) = 0$. The Markov generator of this process can be written in the form

$$\mathcal{L}_{\mathcal{M}_z} = \{\cdot, H\}_{\mathcal{M}_z} + \frac{1}{\beta} \left\{ \mathrm{d}, \left\{ \mathrm{d}^T, \cdot \right\}_{\mathcal{M}_z} e^{-\beta H} \right\}_{\mathcal{M}_z} e^{\beta H},$$

demonstrating the reversibility with respect to the constrained canonical measure $e^{-\beta H} \sigma_{T^* \mathcal{M}_z}(dp dq)$.

Appendix C. Exact sampling of fluctuation/dissipation perturbations. In this section, we recall how to perform exact sampling of fluctuation/dissipation perturbations. Since we only work with impulses, we refer to the system by using the impulse variables p only. Note that throughout the paper, we also use extended variables (p, p_z) , however, the presentation that follows covers general cases. The kinetic energy of the system is $\frac{1}{2} p^T M p$. We impose constraints $p^T M^{-1} \nabla_q \xi = 0$ on impulses, thus $p \in T_q^* \mathcal{M}$ and hence the associated orthogonal projector on $T_q^* \mathcal{M}$ is

$$P = \mathrm{Id} - \nabla_q \xi G^{-1} \nabla_q^T \xi M^{-1}.$$

The stochastic differential equations of motion on impulses that are integrated on a time-step interval are

$$\begin{cases} \dot{p} = -\gamma M^{-1} p + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ p^T M^{-1} \nabla_q \xi = 0, \end{cases} \quad (C_p) \quad (C.1)$$

with the usual fluctuation/dissipation relation $\sigma \sigma^T = 2\beta^{-1} \gamma$. The Gaussian distribution of impulses

$$\frac{1}{Z} e^{-\frac{\beta}{2} p^T M^{-1} p} \sigma_{T_q^* \mathcal{M}}(dp) \quad (C.2)$$

is invariant under the dynamics (C.1).

PROPOSITION C.1 (Exact sampling of stochastic perturbation). *Given the mass matrix M , suppose either δt or γ are small enough so that the condition*

$$\frac{\delta t}{2} M^{-1} \leq \gamma \quad (C.3)$$

holds in the sense of symmetric semi-definite matrices. Let U be a centered and normalized Gaussian vector. Consider the mid-point Euler scheme with constraints

$$\begin{cases} p_{n+1} = p_n - \frac{\delta t}{2} \gamma M^{-1} (p_n + p_{n+1}) + \sqrt{\delta t} \sigma U - \nabla_q \xi \lambda_{n+1} \\ p_{n+1}^T M^{-1} \nabla_q \xi = 0, \end{cases} \quad (C_p) \quad (C.4)$$

where λ_{n+1} is the Lagrange multiplier associated with the constraint (C_p). The Markov kernel defined by the transition $p_n \rightarrow p_{n+1}$ is reversible with respect to the Gaussian distribution (C.2).

Proof. After calculating the Lagrange multiplier the expression (C.4) can be written as

$$p_{n+1} = p_n - \frac{\delta t}{2} P \gamma P^T M^{-1} (p_n + p_{n+1}) + \sqrt{\delta t} P \sigma U.$$

Consider the new variable $\tilde{p} = \beta^{1/2} M^{-1/2} p$, and define the symmetric matrix

$$L \equiv \frac{\delta t}{2} M^{-1/2} P \gamma P^T M^{-1/2},$$

as well as K , such that $KK^T = L$. In terms of these new variables we obtain from (C.4)

$$\tilde{p}_{n+1} = (\text{Id} + L)^{-1}(\text{Id} - L)\tilde{p}_n + 2(\text{Id} + L)^{-1}K U. \quad (\text{C.5})$$

Moreover, the product measure $\sigma_{T_q^* \mathcal{M}_0}(dp_n) \sigma_{T_q^* \mathcal{M}_0}(dp_{n+1})$ is the measure induced on the linear subspace of constraints by the scalar product M^{-1} and the Lebesgue measure $dp_n dp_{n+1}$. Thus in the variables $(\tilde{p}_n, \tilde{p}_{n+1})$ this measure becomes, up to a constant, the measure induced by the usual Euclidean structure. As a consequence the log density of the random variable $(\tilde{p}_n, \tilde{p}_{n+1})$ defined by (C.5) with respect to this latter measure is equal to

$$\begin{aligned} & -\frac{1}{2} |\tilde{p}_n|^2 - \frac{1}{8} (\tilde{p}_{n+1} - (\text{Id} + L)^{-1}(\text{Id} - L)\tilde{p}_n)^T L^{-1}(\text{Id} + L)^2 (\tilde{p}_{n+1} - (\text{Id} + L)^{-1}(\text{Id} - L)\tilde{p}_n) \\ & = -\frac{1}{8} \tilde{p}_{n+1}^T L^{-1}(\text{Id} + L)^2 \tilde{p}_{n+1} - \frac{1}{8} \tilde{p}_n^T L^{-1}(\text{Id} + L)^2 \tilde{p}_n, \end{aligned}$$

which is indeed symmetric between \tilde{p}_n and \tilde{p}_{n+1} . Hence we have shown the reversibility of the induced Markov kernel and consequently stationarity of the canonical Gaussian distribution. \square

REFERENCES

- [1] V. I. Arnol'd. *Mathematical methods of classical mechanics*. Springer-Verlag, New York, 1989.
- [2] M. Beccaria and G. Curci. The Kramers equation simulation algorithm: 1. Operator analysis. *Phys. Rev. D*, 49:2578–2589, 1994.
- [3] C.H. Bennett. Mass tensor molecular dynamics. *J. Comp. Phys.*, 19:267–279, 1975.
- [4] A. Beskos, G.O. Roberts, A.M. Stuart, and J. Voss. An MCMC method for diffusion bridges. *Pre-print*, 2007.
- [5] F. Bornemann and C. Schuette. Homogenization of Hamiltonian system with a strong constraining potential. *Physica D*, 102:57–77, 1992.
- [6] C. Le Bris and F. Legoll. Derivation of symplectic numerical schemes for highly oscillatory Hamiltonian systems. *C. R. Acad. Sci. Paris*, 344:277–282, 2007.
- [7] E. Cances, F. Legoll, and G. Stoltz. Comparison of NVT sampling methods. Technical Report 2040, IMA, 2005.
- [8] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. Technical Report 309, CERMICS, 2006.
- [9] M. E. Clamp, P. G. Baker, C.J. Stirling, and A. Brass. Hybrid Monte Carlo : an efficient algorithm for condensed matter simulation. *J. Comput. Chem.*, 15(8):838–846, 1994.
- [10] D. Cohen, T. Jahnke, K. Lorenz, and Ch. Lubich. Numerical integrators for highly oscillatory Hamiltonian systems: a review. In Alexander Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*, pages 553–576. Springer, 2006.
- [11] M. Creutz and A. Gocksch. Higher-order hybrid Monte Carlo algorithms. *Phys. Rev. Lett.*, 63(1):9–12, 1988.
- [12] S. Duane. Stochastic quantization versus the microcanonical ensemble - getting the best of both worlds. *Nuclear Physics B*, 257(5):652–662, 1985.
- [13] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett.*, 195B(2):216–222, 1987.
- [14] C. Lubich, E. Hairer, G. Wanner. *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2002.
- [15] G. Ciccotti, E. Vanden-Eijnden. Second-order integrators for Langevin equations with holonomic constraints. *Chem. Phys. Lett.*, 429(1-3):310–316, 2006.
- [16] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.
- [17] H. Federer. *Geometric measure theory*. Springer, New-York, 1969.
- [18] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, New-York, 1971.
- [19] M. Fixman. Simulation of polymer dynamics. I. General theory. *J. Chem. Phys.*, 69:1527, 1978.
- [20] M. Fixman. Implicit algorithm for Brownian dynamics of polymers. *Macromolecules*, 19:1195–1204, 1986.
- [21] E. Vanden-Eijnden, G. Ciccotti, R. Kapral. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *J. Chem. Phys.*, 6(9):1809–14, 2005.

- [22] A. M. Horowitz. A generalized guided Monte Carlo algorithm. *Phys. Lett.B*, 268:247–252, 1991.
- [23] I. Horváth and A. D. Kennedy. The Local Hybrid Monte Carlo algorithm for free field theory. *Nucl. Phys. B*, 510:367–400, 1998.
- [24] J. A. Izaguirre and S. S. Hampton. Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules. *J. Comput. Phys.*, 200(2):581–604, 2004.
- [25] H.J.C. Berendsen J-P. Ryckaert; G. Ciccotti. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-Alkanes. *J. of Comp. Phys.*, 23:327–341, 1977.
- [26] H.J.C. Berendsen K. A. Feenstra, B. Hess. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. of Comp. Chemistry*, 20:786–798, 1999.
- [27] N.G. Van Kampen. Elimination of fast variables. *Physics Reports*, 124(2):9–160, 1985.
- [28] B. J. Leimkuhler and R. D. Skeel. Symplectic numerical integrators in constrained Hamiltonian systems. *J. Comput. Phys.*, 112(1):117–125, 1994.
- [29] B. Mao and A.R. Friedman. Molecular dynamics simulation by atomic mass weighting. *Biophysical Journal*, 58:803–805, 1990.
- [30] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machine. *J. Chem. Phys.*, 21:1087–1091, 1953.
- [31] B. Oksendal. *Stochastic differential equations (3rd ed.): an introduction with applications*. Springer-Verlag, 1992.
- [32] S. Reich. Smoothed dynamics of highly oscillatory Hamiltonian systems. *Physica D*, 89:28–42, 1995.
- [33] S. Reich. Smoothed Langevin dynamics of highly oscillatory systems. *Physica D*, 138:210–224, 2000.
- [34] S. Reich and B. Leimkuhler. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2005.
- [35] H. Rubin and P. Ungar. Motion under a strong constraining force. *Comm. Pure Appl. Math*, 10:65–87, 1957.
- [36] Ch. Schütte, A. Fischer, W. Huisinga, and P. Deuffhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys.*, 151(1):146–168, 1999.
- [37] F. Takens. Motion under the influence of a strong constraining force. In *Global theory of dynamical systems (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979)*, volume 819 of *Lecture Notes in Math.*, pages 425–445. Springer, Berlin, 1980.
- [38] W.F. van Gunsteren and H. J. C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Molecular Physics*, 34(5):1311–1327, 1977.
- [39] W.F. van Gunsteren and H.J.C. Berendsen. Algorithms for Brownian dynamics. *Molecular Physics*, 45:637–647, 1982.